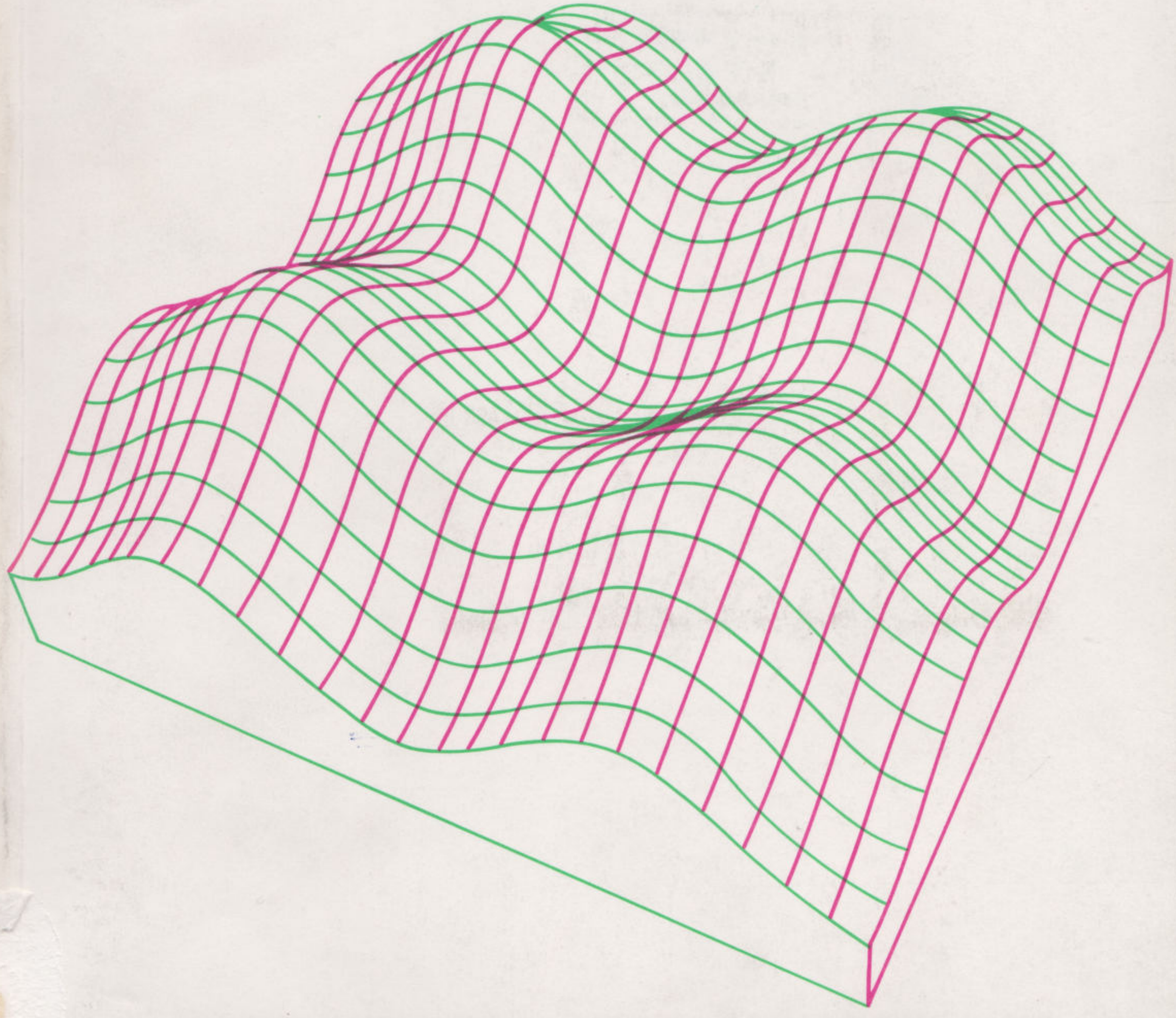
M 100 15

THE OPEN UNIVERSITY

Mathematics Foundation Course Unit 15

Differentiation II



The Open University

Mathematics Foundation Course Unit 15

DIFFERENTIATION II

Prepared by the Mathematics Foundation Course Team

Correspondence Text 15

The Open University Press

Open University courses provide a method of study for independent learners through an integrated teaching system including textual material, radio and television programmes and short residential courses. This text is one of a series that make up the correspondence element of the Mathematics Foundation Course.

The Open University's courses represent a new system of university level education. Much of the teaching material is still in a developmental stage. Courses and course materials are, therefore, kept continually under revision. It is intended to issue regular up-dating notes as and when the need arises, and new editions will be brought out when necessary.

Further information on Open University courses may be obtained from The Admissions Office, The Open University, P.O. Box 48, Bletchley, Buckinghamshire.

The Open University Press
Walton Hall, Bletchley, Bucks.

First published 1971

Printed in Great Britain by
J W Arrowsmith Ltd, Bristol 3

SBN 335 01014 8

Contents

Objectives

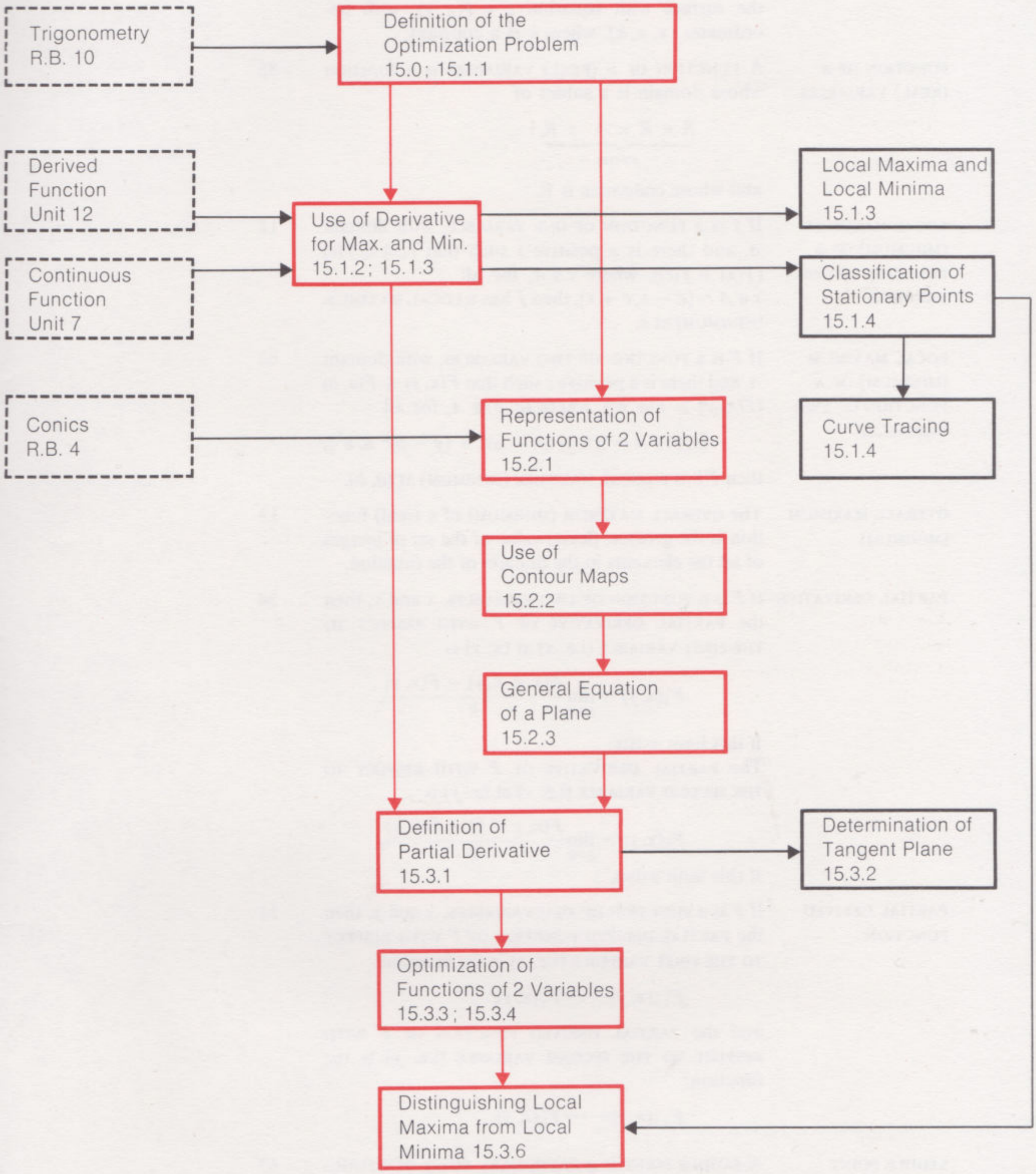
After working through this unit you should be able to:

(i) locate the local maxima and minima of a function of one variable;
(ii) determine the greatest or least value of the images of a function of one variable;
(iii) determine the behaviour of a function of one variable near the zeros of its derived function;
(iv) sketch a rough graph of a given (simple) function;
(v) express the equations of simple surfaces using Cartesian co-ordinates in three dimensions;
(vi) represent a function of two variables as a three-dimensional surface;
(vii) define the partial derivatives of a function of two variables, and evaluate these derivatives in simple cases;
(viii) determine the equation of a tangent plane to a given surface at a given point;
(ix) determine the local maxima and minima of a function of two variables;
(x) distinguish between local maxima and minima of a function of two variables, x and y, by determining the behaviour of curves of intersection between the surface and planes through the point P, where the tangent plane to the surface at P is parallel to the xy-plane.

Note

Before working through this correspondence text, make sure you have read the general introduction to the mathematics course in the Study Guide, as this explains the philosophy underlying the whole course. You should also be familiar with the section which explains how a text is constructed and the meanings attached to the stars and other symbols in the margin, as this will help you to find your way through the text.

Structural Diagram



Glossary

Page

Terms which are defined in this glossary are printed in CAPITALS.

CONTOUR LINE — A CONTOUR LINE is a line on a SURFACE marking a particular level; for example, a subset of points on the surface with equation $z = F(x, y)$, with co-ordinates (x, y, k), where k is a constant. 45

FUNCTION OF n (REAL) VARIABLES — A FUNCTION OF n (REAL) VARIABLES is a function whose domain is a subset of

$$\underbrace{R \times R \times \cdots \times R}_{n\,\text{times}},$$

and whose codomain is R. 35

LOCAL MAXIMUM (MINIMUM) OF A FUNCTION OF ONE VARIABLE — If f is a FUNCTION OF ONE VARIABLE, with domain A, and there is a positive ε such that $f(x) \leqslant f(c)$ $(f(x) \geqslant f(c))$, where $c \in A$, for all $x \in A \cap [c - \varepsilon, c + \varepsilon]$, then f has a LOCAL MAXIMUM (MINIMUM) at c. 12

LOCAL MAXIMUM (MINIMUM) OF A FUNCTION OF TWO VARIABLES — If F is a FUNCTION OF TWO VARIABLES, with domain A, and there is a positive ε such that $F(x, y) \leqslant F(a, b)$ $(F(x, y) \geqslant F(a, b))$, where $(a, b) \in A$, for all

$$(x, y) \in A \cap \{(x, y) : (x - a)^2 + (y - b)^2 \leqslant \varepsilon^2\},$$

then F has a LOCAL MAXIMUM (MINIMUM) at (a, b). 62

OVERALL MAXIMUM (MINIMUM) — The OVERALL MAXIMUM (MINIMUM) of a (real) function is the greatest (least) value of the set of images of all the elements in the domain of the function. 12

PARTIAL DERIVATIVE — If F is a FUNCTION OF TWO VARIABLES, x and y, then the PARTIAL DERIVATIVE OF F WITH RESPECT TO THE FIRST VARIABLE (i.e. x) at (x, y) is 54

$$F'_1(x, y) = \lim_{h \to 0} \frac{F(x + h, y) - F(x, y)}{h},$$

if this limit exists.
The PARTIAL DERIVATIVE OF F WITH RESPECT TO THE SECOND VARIABLE (i.e. y) at (x, y) is

$$F'_2(x, y) = \lim_{k \to 0} \frac{F(x, y + k) - F(x, y)}{k},$$

if this limit exists.

PARTIAL DERIVED FUNCTION — If F is a FUNCTION OF TWO VARIABLES, x and y, then the PARTIAL DERIVED FUNCTION OF F WITH RESPECT TO THE FIRST VARIABLE (i.e. x) is the function: 54

$$F'_1 : (x, y) \longmapsto F'_1(x, y),$$

and the PARTIAL DERIVED FUNCTION OF F WITH RESPECT TO THE SECOND VARIABLE (i.e. y) is the function:

$$F'_2 : (x, y) \longmapsto F'_2(x, y).$$

SADDLE POINT — A SADDLE POINT is a STATIONARY POINT of a FUNCTION OF TWO VARIABLES which is neither a LOCAL MAXIMUM nor a LOCAL MINIMUM. 63

STATIONARY POINT OF A FUNCTION OF ONE VARIABLE — A STATIONARY POINT OF A FUNCTION f OF ONE VARIABLE, with domain A, is a point $c \in A$ such that 15

$$f'(c) = 0.$$

		Page
STATIONARY POINT OF A FUNCTION OF TWO VARIABLES	A STATIONARY POINT OF A FUNCTION F OF TWO VARIABLES, with domain A, is a point $(a, b) \in A$ such that $F_1'(a, b) = F_2'(a, b) = 0$.	61
SURFACE	A SURFACE is the geometric figure consisting of those points whose co-ordinates, x, y, z, satisfy any equation such as $z = f(x, y)$ or $F(x, y, z) = 0$. Alternatively, x, y and z may be given by three *parametric* equations of the form: $x = G(r, \alpha), y = H(r, \alpha)$, $z = K(r, \alpha)$ where G, H and K are FUNCTIONS OF TWO VARIABLES, r and α (called *parameters*).	35
TANGENT PLANE	The TANGENT PLANE to a SURFACE at a point P (on the surface) is the plane which is such that each line in the plane which passes through P is a tangent to the surface at P.	57

Notation

Page

The symbols are presented in the order in which they appear in the text.

Bibliography

For a light introduction to three-dimensional geometry and its place in the general pattern of mathematics, see

W. W. Sawyer, *A Path to Modern Mathematics*, (Penguin Books, 1966).

For a more detailed treatment of functions of one and two variables, which is similar to our own, see

T. M. Apostol, *Calculus Vol. I*, (Blaisdell, 1967).

For applications to physical problems of the techniques discussed in this unit, and discussion of more advanced techniques see

Ben Noble, *Applications of Undergraduate Mathematics in Engineering*, (Collier-Macmillan, 1967).

15.0 INTRODUCTION

15.0

Introduction

Many problems in both pure and applied mathematics are concerned with maximum or minimum properties of some sort. For example, at what angle should a missile be fired in order to give the maximum range? What is the largest area which can be surrounded by a given length of fencing? What is the shortest path between two points on a given surface? Problems of this kind are sometimes called *optimization problems*, and some of them can be attacked systematically using calculus.

In this unit we are mainly concerned with the problem of determining the greatest and least values attained by the images of a given function, and the elements in the domain to which these images correspond. Such problems occur quite frequently in practical situations, but, as one would expect, the functions which arise in realistic cases tend to be rather formidable. For this reason the "practical" situations which we discuss here are grossly simplified, and, as in our first example, we sometimes make no pretence of realism at all. However, the techniques which we develop are often used to discuss non-trivial systems, and it is the wide variety of applications which makes these methods so useful.

We have already seen some methods for solving problems of this kind in *Unit 6, Inequalities*, but the discussion there applied mainly to linear functions: here the functions are more general.

Essentially, the aim of this unit is to bring to your attention some techniques and ideas which can be developed into rigorous methods for the study of optimization, and to point out those areas which will need a deeper treatment in later years.

We begin by discussing functions of one real variable, by which we mean *real functions* (whose domains and codomains are R or subsets of R).

The second part of the text is devoted to a similar treatment for real-valued functions of two real variables (functions which map $R \times R$, or subsets thereof, to R). Before tackling this we discuss a little three-dimensional co-ordinate geometry which we need to know, and which, in any case, is of interest in its own right.

You will probably have most difficulty with sections 15.2 and 15.3, mainly because these sections deal with three-dimensional situations. For this reason the television programme is devoted almost entirely to the geometric notions which we shall need in three dimensions: the ideas of planes, surfaces and so on.

15.1 OPTIMIZATION OF FUNCTIONS OF ONE VARIABLE

15.1

15.1.1 A Cautionary Tale

15.1.1

Discussion

A mathematically minded Chancellor of the Exchequer once had what he thought was a wonderful idea for a new sort of tax. He would levy a tax on the amount of overtime that people worked. The idea seemed to be superb from every point of view. For their own good he would be deterring people from working long hours, and it was merely incidental that a large sum would be raised annually for the Exchequer. Thus his humanitarian principles and the requirements of his office would be satisfied together.

This is what he intended to do. For the first hour worked in excess of forty hours, the worker would be required to pay 5p, for the next hour 10p, and so on. What could be easier?

A fellow called Fred earned 50p an hour in a factory, and he was puzzled. How many hours should he work in order to earn the most money?

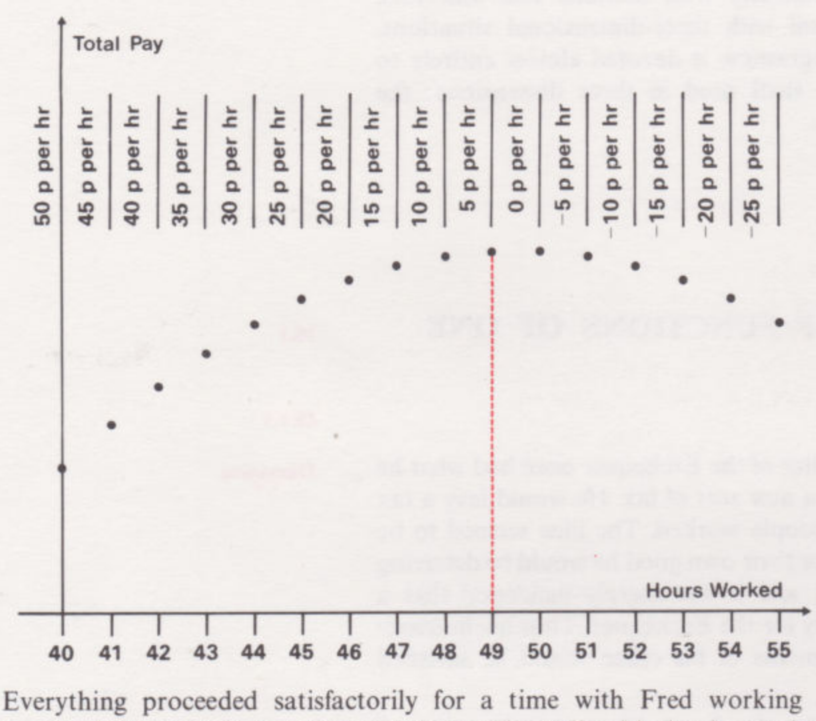
For 40 hours work he would get £20.00;
for 41 hours work he would get £20.45;
for 42 hours work he would get £20.85;

and so on.

After calculating the total pay for all the possible numbers of working hours from 40 to 55, he could see that it was a waste of time to work for more than 49 hours per week.

Hours worked	Total pay (£)
40	20.00
41	20.45
42	20.85
43	21.20
44	21.50
45	21.75
46	21.95
47	22.10
48	22.20
49	22.25
50	22.25
51	22.20
52	22.10
53	21.95
54	21.75
55	21.50

His wife looked over his shoulder (as wives will) as he worked all this out, and promptly told him he was a fool. *Obviously* he should not work for more than 49 hours, because the *rate* at which he earned money went down from 50p an hour in the 40th hour by steps of 5p an hour until in the 50th hour the rate was down to zero, after which his rate of earning was negative and he would simply be giving money away.



Everything proceeded satisfactorily for a time with Fred working a regular 49 hour week, but after a while the Chancellor felt that it was time to add a few refinements to keep the tax inspectors on their toes, and to keep his name before the great voting public.

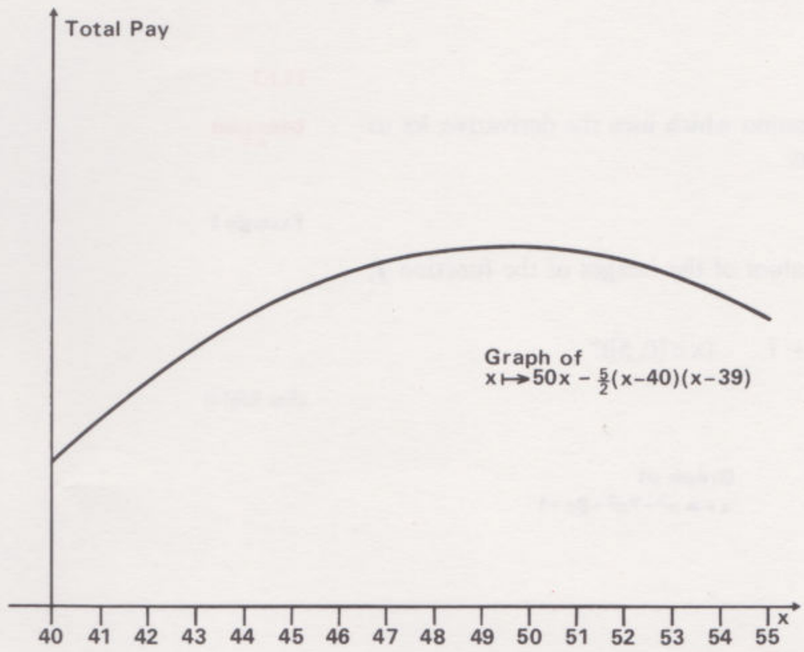
What if someone were to work for some fraction of an hour? The Chancellor had been keenly following the Open University's Mathematics Foundation Course, and was anxious to try out his new-found knowledge of finite differences. It was quite a simple matter for him to formulate the rule for calculation of "Overtime Tax":

"The tax deducted shall be $\frac{5}{2}(x - 40)(x - 39)$ new pence, where x is the total number of hours (or fraction thereof) worked in one week."

He didn't find it difficult to verify that the formula gave him the same result as before when x was any whole number greater than 40.

Fred then had the problem of finding the value of x which gave him the maximum pay. He was fairly certain that it was between 49 and 50, but to determine it exactly was simply beyond him. All he could say was that his pay would be:

$$50x - \tfrac{5}{2}(x - 40)(x - 39) \text{ new pence per week.}$$



Luckily Fred never had to solve the problem, for the very next day the Chancellor caught a cold and could not go to work for a week. At the end of that time he was presented with a bill for £39.00 from the local tax office. He promptly changed the law.

This story brings out some of the important points of the first section of this text, in which we wish to optimize functions of one variable. Before the introduction of the Chancellor's amendment, it might seem to be a sound practical technique to calculate the total pay for each possible number of hours worked, and then choose the largest value. Fred chose to do this from 40 to 55 hours, but why stop there? He could be sure that his pay wouldn't be more than £22.25 after 73 hours, say, but what would he have done if the tax law had been more complicated? The method of looking at the image of every element in the domain of the function under consideration (in this case the function which maps hours worked to total pay) will only work when the domain is finite. Even then, it may involve a considerable amount of calculation if the domain has a large number of elements.

When the domain of the function is not a finite set, such as the interval of real numbers in our example, then we must think of some other technique. The answer lies in a comment from Fred's wife: "Stop working once your *rate* of earning is zero, because then you have reached your maximum

wage." There is a little more to the technique we shall develop, but essentially that is the idea. As another example on the same line of thought, you reached your maximum height once your *rate* of growth became zero, in other words, when you stopped growing.

In this first section, then, we shall use the concept of *rate of change* to solve optimization problems for functions of one variable. Clearly we are going to need some results concerning rate of change from *Unit 12, Differentiation I.*

Exercise 1

Exercise 1
(5 minutes)

A supermarket sells chocolate biscuits at 20p a packet and sells 1200 packets weekly. The manager estimates that for every penny by which the price is reduced, sales will increase by 300 packets a week. The lowest possible price is 12p, which is what the supermarket pays the wholesaler. At what price should the manager sell the biscuits in order to maximize his profit? ■

15.1.2 Using the Derivative

15.1.2

Discussion
* *

To introduce a method of optimization which uses the derivative, let us first look at a fairly simple function.

Example 1

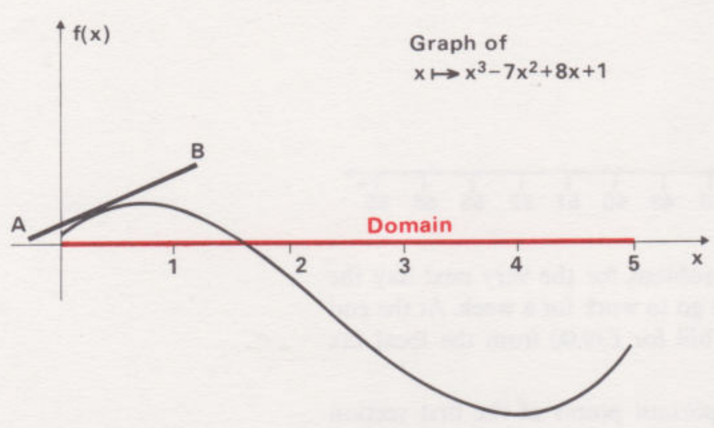
Example 1

What are the greatest and least values of the images of the function f, where

$$f: x \longmapsto x^3 - 7x^2 + 8x + 1 \qquad (x \in [0, 5])?$$

First we look at the graph of f:

(See RB14)



Imagine the tangent line AB moving along the curve from the point where $x = 0$ to the point where $x = 5$. The slope of this line is initially positive, becomes negative, and is positive again when we reach $x = 5$. At two intermediate points the line is parallel to the x-axis (it has zero slope) and the graph shows that these are the points at which $f(x)$ takes its greatest and least values in the interval $[0, 5]$. Remember that $f'(x)$ is the slope of the tangent at x. If

$$f(x) = x^3 - 7x^2 + 8x + 1 \qquad (x \in [0, 5]),$$

then we know from *Unit 12, Differentiation I* that the slope at x is given by

$$f'(x) = 3x^2 - 14x + 8.$$

The values of x for which $f'(x) = 0$ are the two solutions of the quadratic equation:

(See RB5)

$$3x^2 - 14x + 8 = 0,$$

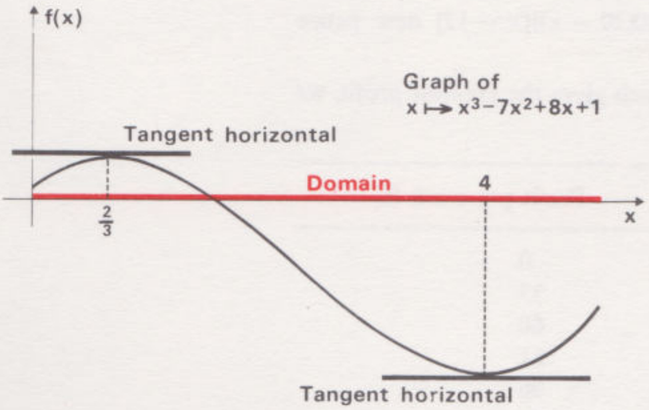
which are

$$x = \tfrac{2}{3} \quad \text{and} \quad x = 4.$$

The greatest and least values of $f(x)$ in the interval $[0, 5]$ are therefore

$$f(\tfrac{2}{3}) = 3\tfrac{14}{27} \quad \text{and} \quad f(4) = -15$$

respectively.



■

Exercise 1

Exercise 1
(4 minutes)

(i) Find the greatest value of $g(x)$, where

$$g: x \longmapsto 4 - x^2 \qquad (x \in [-2, 2]).$$

(ii) Find the greatest rectangular area which can be enclosed by a fence of length 100 metres. ■

Discussion

This technique has to be used with a certain amount of care on some occasions. The following examples illustrate the sort of difficulties which can occur.

Example 2

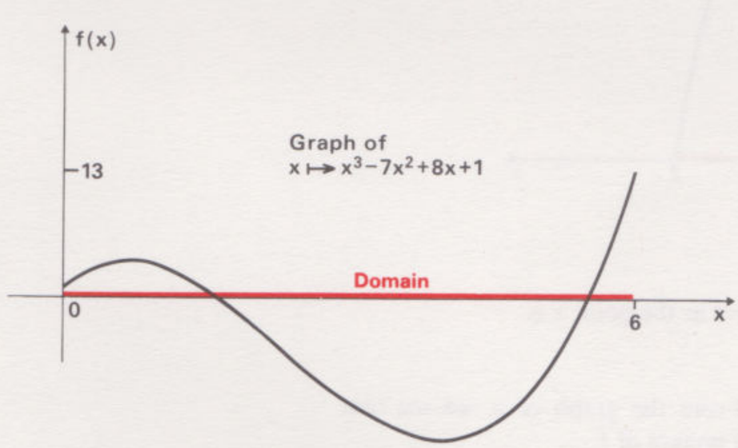
Example 2

What is the greatest value of

$$f(x) = x^3 - 7x^2 + 8x + 1 \qquad \text{in } [0, 6]?$$

You may well say that the answer is $3\frac{14}{27}$, as in Example 1. But, if so, how do you explain the fact that $f(6) = 13$?

The apparent contradiction is explained when we examine the graph of the function, which shows that the greatest value of $f(x)$ in $[0, 6]$ occurs when $x = 6$.



■

(continued on page 7)

Solution 15.1.1.1

Solution 15.1.1.1

Suppose that the supermarket manager fixed a price of x new pence; his profit on each packet would then be $(x - 12)$ new pence. He would have reduced his sales price by $(20 - x)$ new pence, and therefore increased his sales by 300 times $(20 - x)$ packets per week. The total number of packets sold per week would be

$$1200 + 300(20 - x),$$

so that his profit would be $(1200 + 300(20 - x))(x - 12)$ new pence per week.

In order to determine the selling price which gives the greatest profit, we need only construct the following table:

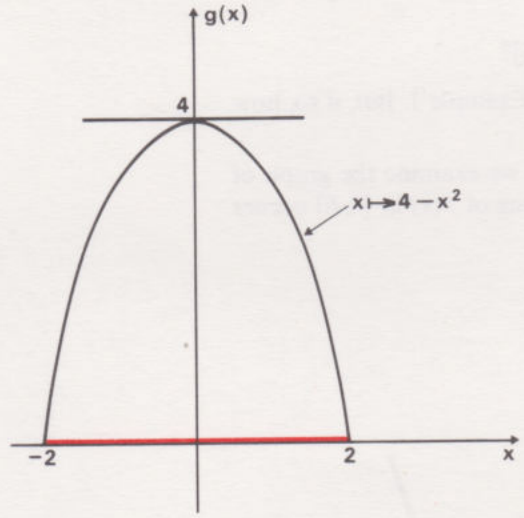
Selling price, x (p)		Profit per week (£)	
12		0	
13		33	
14		60	
15		81	
16		96	
17		105	
18	optimum selling price	108	greatest profit
19		105	
20		96	

There are other ways of solving this problem, as we shall see later. ■

Solution 1

Solution 1

(i) Since $x^2 \geqslant 0$ for all x, the greatest image is $g(0) = 4$. To illustrate the method we are introducing, we also sketch the graph of g:

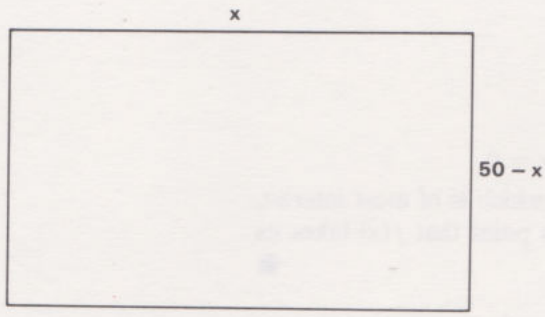


The slope of the tangent to the curve at the point x is

$$g'(x) = -2x$$

This slope is zero when $x = 0$. From the graph of g, we see that $g(0) = 4$ is the greatest of the set of images of f.

(ii)



Let x be the length in metres of one side of the rectangle. The area of the rectangle is

$$x(50 - x)\,\text{m}^2,$$

so we can express the area of the rectangle by the function:

$$g: x \longmapsto x(50 - x) \qquad (x \in [0, 50]).$$

Then

$$g'(x) = 50 - 2x$$

and

$$g'(x) = 0 \text{ when } x = 25.$$

With this value for x, the rectangle is a square, and the required area is $625\,\text{m}^2$. ■

(*continued from page 5*)

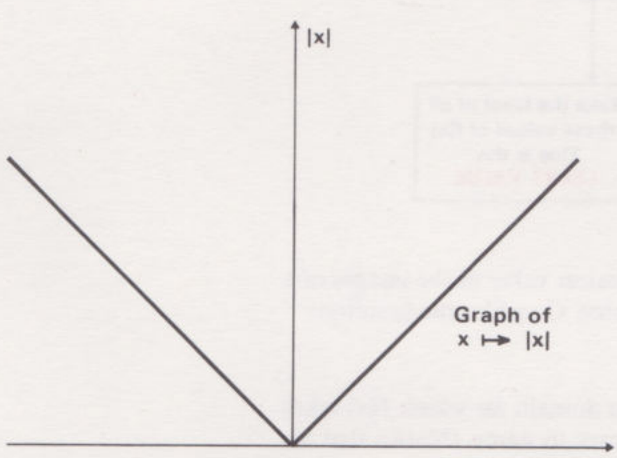
Example 3

Example 3

What is the least value of $f(x)$, where

$$f: x \longmapsto |x| \qquad (x \in R)?$$

This function (which you first met in *Unit 1, Functions*) is called the *modulus function*; it has the following graph:



The difficulty in this case is that f is not differentiable for all values of x. We saw in section 12.2.1 of *Unit 12, Differentiation I* that

$$f'(x) = +1 \qquad \text{if } x > 0,$$

and

$$f'(x) = -1 \qquad \text{if } x < 0;$$

but if $x = 0$, then the limit:

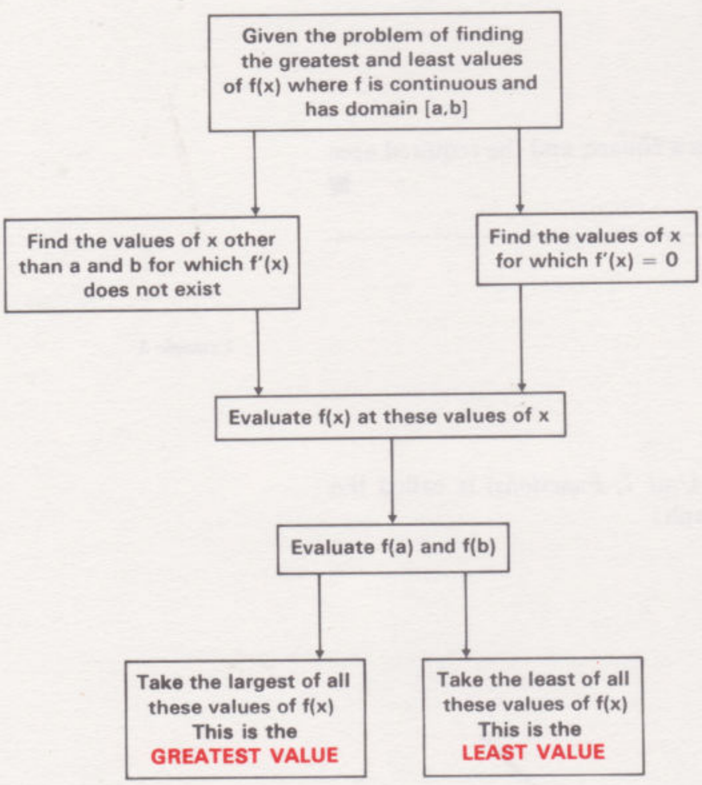
$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

does not exist, and so f is not differentiable at $x = 0$.

Unfortunately in this case it is this very point which is of most interest, for we can see from the graph that it is at this point that $f(x)$ takes its least value. ■

The following strategy will enable us to deal with most problems (and certainly all the problems of this type which occur in the Foundation Course) without having to draw graphs. (It is nearly always useful, however, to draw a rough sketch of the graph when this is not too difficult.) There are cases in which things can still go wrong, but they will not concern us here.

A Strategy for Finding the Greatest and Least Values of the Images of a Well-behaved* Function



Sometimes it is impossible to specify the greatest value of the images of a function, simply because there is no such value. Consider the function:

$$f : x \longmapsto x^2 \qquad (x \in R).$$

In this case, there are always elements in the domain for which $f(x)$ takes values greater than any fixed number you care to name. (Notice that we may not say "the greatest value is infinity" because infinity is *not* a number, and by the words "greatest value" we imply that we are looking for a number.)

* By a "well-behaved" function we mean a continuous function which is differentiable at all but a *finite* number of points in its domain. It *is* possible to find a function which is continuous at all points in its domain, but not differentiable at *any* point in its domain, and to which the above strategy does not apply. (For the definition of a continuous function, see *Unit 7, Sequences and Limits I*.)

Exercise 2

Exercise 2
(5 minutes)

Find the greatest and least values (whenever possible) of the images of the following functions, and sketch the graph of each function.

(See RB14)

(i) $h : x \longmapsto 300(36x - 288 - x^2) \qquad (x \in [12, 24])$.
(Compare this with Exercise 15.1.1.1.)

(ii) $f : x \longmapsto x + \dfrac{1}{x} \qquad (x \in R^+)$

(iii) $g : x \longmapsto x^2 + \dfrac{16}{x} \qquad (x \in R^+)$.

■

Solution 2 **Solution 2**

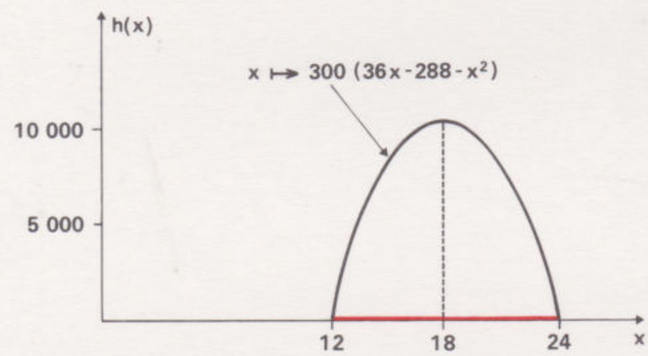
(i) You will find that the supermarket manager's profit function (Exercise 15.1.1.1) can be rearranged to give the same rule as the function which we are discussing here:

$$h: x \longmapsto 300(36x - 288 - x^2) \qquad (x \in [12, 24]).$$

(The domains of the two functions are different.) We have

$$h'(x) = 300(36 - 2x)$$

so that $h'(x) = 0$ when $x = 18$. This gives the greatest value of $h(x)$, which is 10 800p. The least value of $h(x)$ in the given domain is 0, and this value occurs at *both* endpoints of the domain, that is, at $x = 12$ and $x = 24$.



(ii) $f(x) = x + \dfrac{1}{x}$,

so that

$$f'(x) = 1 - \frac{1}{x^2}$$

and therefore

$$f'(x) = 0 \quad \text{if} \quad x = \pm 1.$$

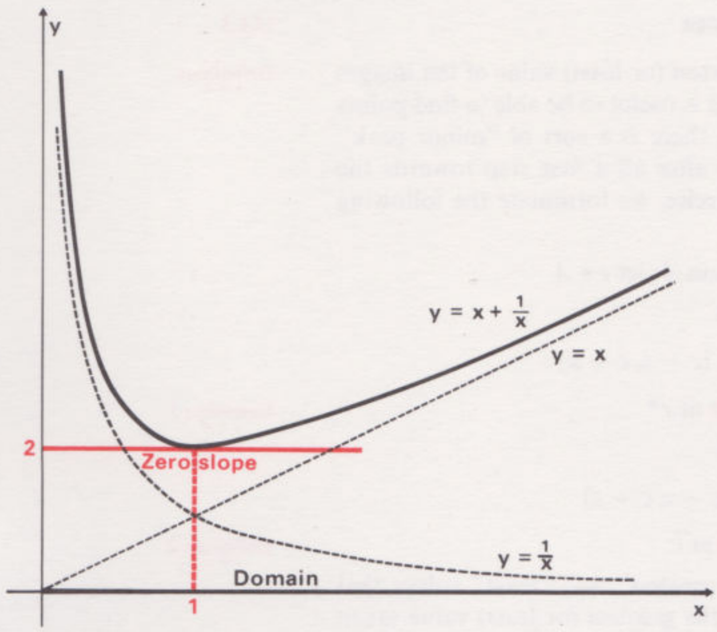
However, only the value $x = 1$ is in the domain of f. On this occasion it helps to sketch a graph of f, and we can easily do this if we notice the following:

(a) $x + \dfrac{1}{x}$ is never zero for $x \in R^+$;

(b) $x + \dfrac{1}{x}$ is very nearly equal to x when x is large, because the value of $\dfrac{1}{x}$ is then very small;

(c) $x + \dfrac{1}{x}$ is very nearly equal to $\dfrac{1}{x}$ when x is small, because then $\dfrac{1}{x}$ is large and predominates over x.

Using the above facts we can draw the following sketch:



The least value of $f(x)$ is $f(1) = 2$; there is no greatest value.

(iii) $g(x) = x^2 + \dfrac{16}{x}$,

so that

$g'(x) = 2x - \dfrac{16}{x^2}$,

and therefore

$g'(x) = 0$ if $2x^3 - 16 = 0$, that is, if $x = 2$.

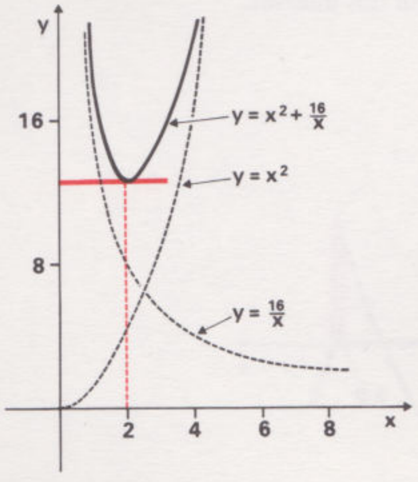
Once again we can draw a sketch if we notice that:

(a) $x^2 + \dfrac{16}{x}$ is a little greater than x^2 when x is very large;

(b) $x^2 + \dfrac{16}{x}$ is a little greater than $\dfrac{1}{x}$ when x is very small (but not zero);

(c) $g(x) > 0$ when $x \in R^+$.

Using the above facts, we can sketch the following graph:



The least value of $g(x)$ is $g(2) = 12$; there is no greatest value. ■

15.1.3 Local Maxima and Minima

15.1.3

Definitions

Although we often wish to find the greatest (or least) value of the images of a function, there are occasions when it is useful to be able to find points (like $x = \frac{2}{3}$ in Example 15.1.2.2) where there is a sort of "minor peak" on the "side of the mountain"; this is after all a first step towards the greatest value. In order to be more precise, we formulate the following definitions.

Let f be a given real function with domain A; let $c \in A$.

If there is a positive number ε such that

$$f(x) \leqslant f(c) \quad \text{for all} \quad x \in A \cap [c - \varepsilon, c + \varepsilon]$$

then we say that f has a local maximum at c*. — Definition 1

If there is a positive number ε such that

$$f(x) \geqslant f(c) \quad \text{for all} \quad x \in A \cap [c - \varepsilon, c + \varepsilon]$$

then we say that f has a local minimum at c. — Definition 2

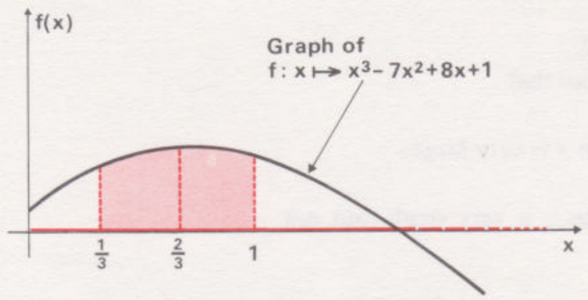
To distinguish these values from the "greatest" and "least" values that we have been discussing, we shall call the greatest (or least) value taken by the images of a function its overall maximum (or minimum). — Definition 3

Discussion

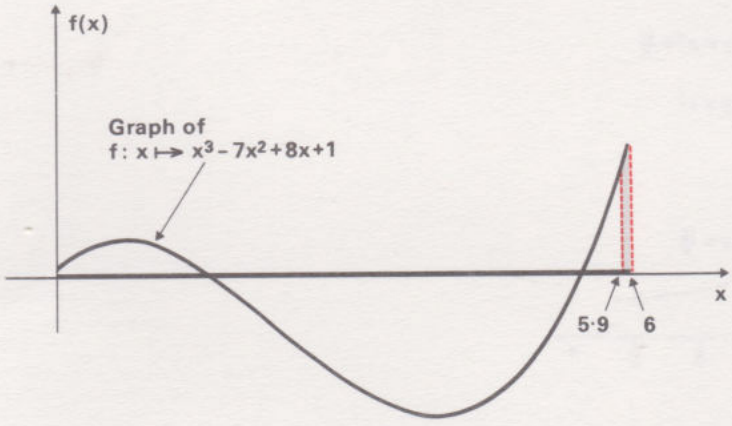
How do the above definitions apply in the context of Example 15.1.2.2? The function:

$$f: x \longmapsto x^3 - 7x^2 + 8x + 1 \qquad (x \in [0, 6])$$

has a local maximum at $\frac{2}{3}$. If we take $\varepsilon = \frac{1}{3}$, say, then $f(x) \leqslant f(\frac{2}{3})$ for all $x \in [\frac{1}{3}, 1]$.



This function also has a local maximum at 6, because for all x in $A = [0, 6]$ close to 6 we have $f(x) \leqslant f(6)$. For example, taking $\varepsilon = 0.1$, the set $A \cap [c - \varepsilon, c + \varepsilon]$ becomes $[5.9, 6]$, and $f(x) \leqslant f(6)$ in this interval.

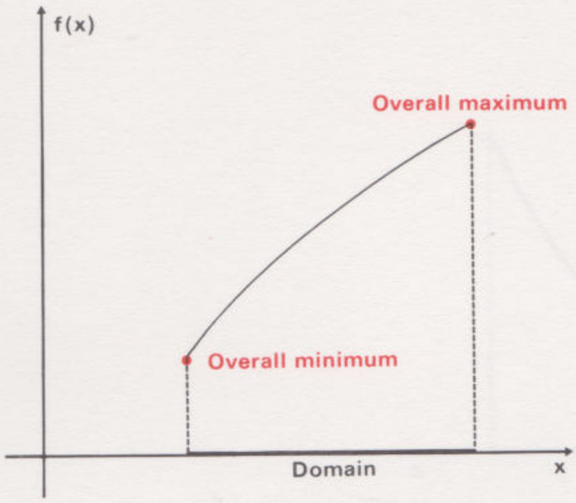


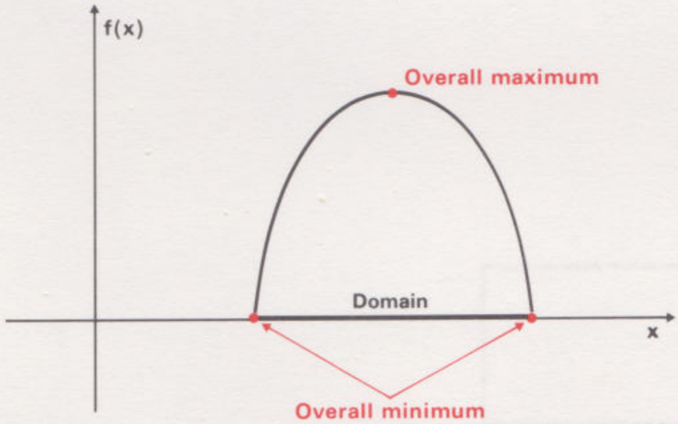
* We shall also say, for example, "$(c, f(c))$ (or simply "c") is a local maximum", when it would be more correct to say "$(c, f(c))$ is a local maximum point".

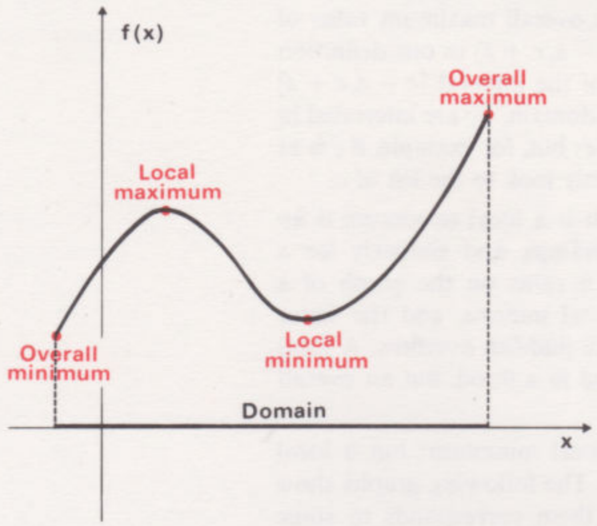
In this case we know that $f(6)$ is in fact the overall maximum value of $f(x)$ for $x \in A$. The reason for taking $A \cap [c - \varepsilon, c + \varepsilon]$ in our definition is that we are only interested in that part of the interval $[c - \varepsilon, c + \varepsilon]$ which lies in A. If c is not an end-point of the domain, we are interested in the behaviour of f close to c *on both sides* of c: but, for example, if c is at the right-hand end of the domain, we need only look to the left of c.

Effectively, we are saying that a point which is a local maximum is an overall maximum in its immediate surroundings, and similarly for a local minimum. Speaking very roughly, if it rains on the graph of a function, the puddles collect around the local minima, and the water runs away to the overall minimum when the puddles overflow. A local maximum would be a suitable place to stand in a flood, but an overall maximum would be preferable.

Of course, an overall minimum is also a local minimum: but a local minimum need not be an overall minimum. The following graphs show some of the various possibilities. Each of them corresponds to some function f.





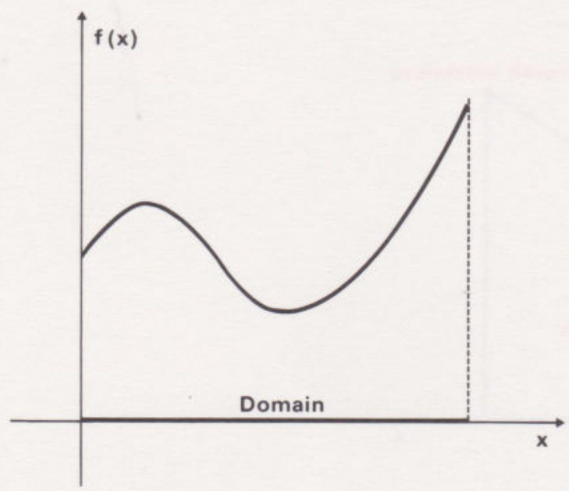


Exercise 1

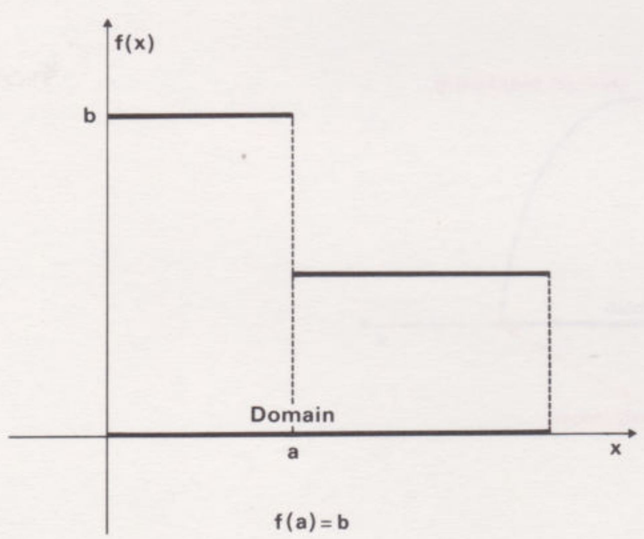
Exercise 1
(5 minutes)

Mark the local and overall maxima and minima on the following graphs:

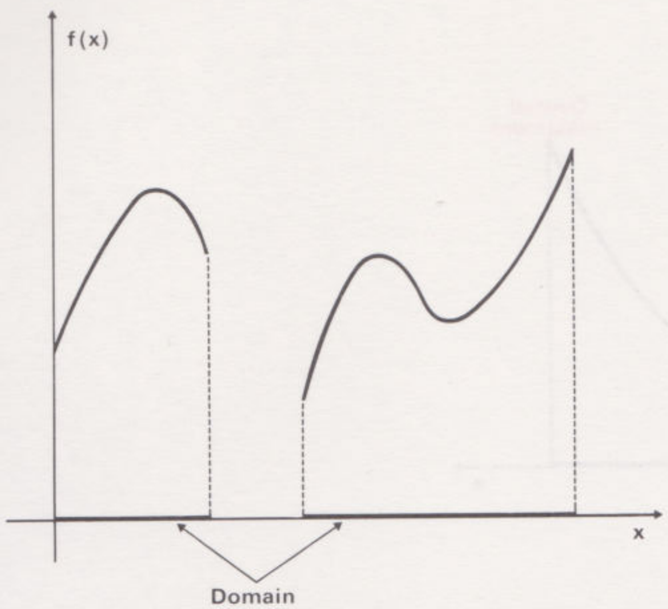
(i)



(ii)



(iii)



■

We call points x such that $f'(x) = 0$ stationary points of f. A stationary point is thus simply a point on the x-axis where the tangent at the corresponding point on the graph is parallel to the x-axis.

Definition 4
* * *

If we wish to locate local maxima (or minima) of a function, it would seem a sound idea to first locate the stationary points. However, there are unfortunately two complications.

Discussion
* *

We have already seen that a local maximum (or minimum) of a function can occur at a point which is not a stationary point (in other words where the slope of the graph is not zero) either because the function is not differentiable at that point, and "slope" is meaningless, or because the point occurs at an end-point of the domain. We shall overcome this difficulty by considering only functions which are differentiable at all points of their domains, and by restricting our search for the local maxima and minima of such a function to points which are not end-points of the domain, and then examining the end-points as a separate issue.

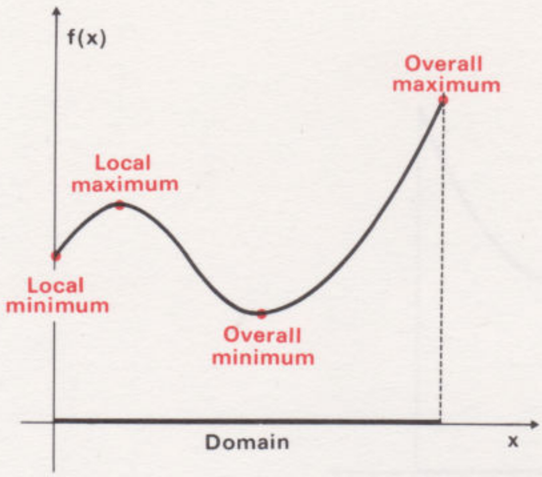
There is a second complication which is more serious: a stationary point may be neither a local maximum nor a local minimum, as we shall see in the next example.

(*continued on page 17*)

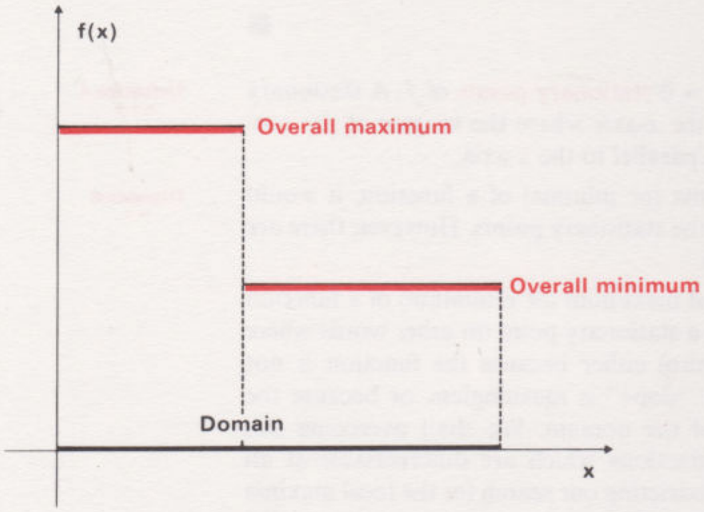
Solution 1

Solution 1

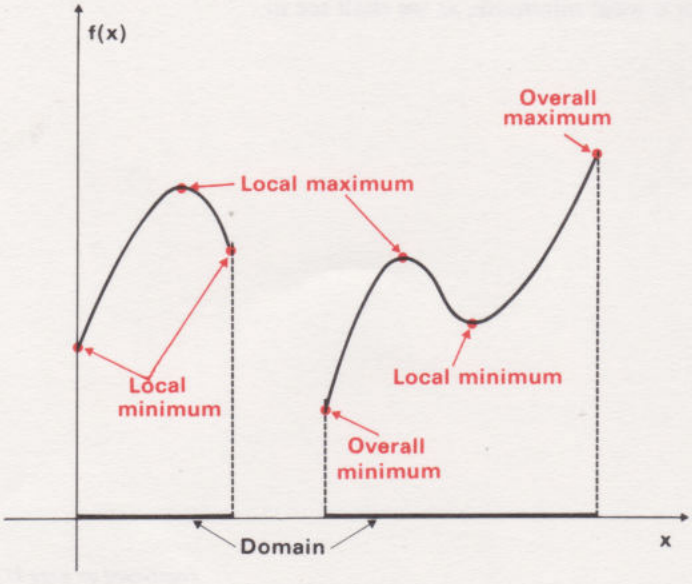
(i)



(ii)



(iii)



(*continued from page 15*)

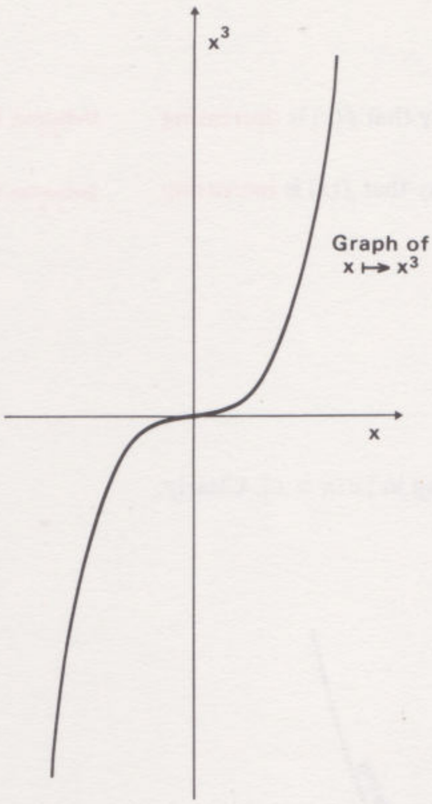
Example 1

Example 1

Consider the function

$$f: x \longmapsto x^3 \qquad (x \in R)$$

which has the graph:



We know from *Unit 12, Differentiation I* that

$$f': x \longmapsto 3x^2 \qquad (x \in R),$$

so that

$$f'(0) = 0,$$

and therefore f has a stationary point at 0. But f has neither a local maximum nor a local minimum at 0. ■

How then can we distinguish between local maxima, local minima and stationary points which are neither?

15.1.4 Two Useful Methods

15.1.4
Technique

In this section we shall describe two methods for determining the nature of stationary points.

Method One

Let f be a real differentiable function:

$$f: x \longmapsto y \qquad (x \in A).$$

Definition 1

If $f'(x) < 0$ for all $x \in S$, where $S \subseteq A$, then we say that $f(x)$ is decreasing in S.

Definition 2

If $f'(x) > 0$ for all $x \in S$, where $S \subseteq A$, then we say that $f(x)$ is increasing in S.

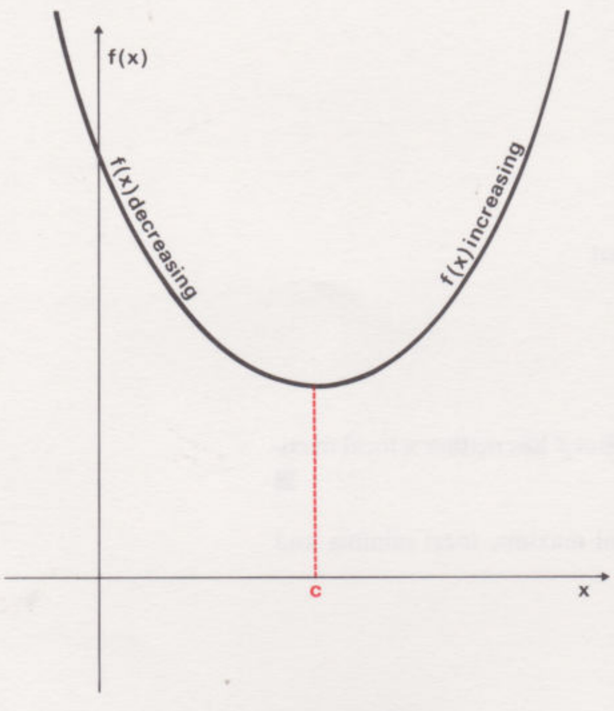
If there is a point $c \in R$ such that

$$f'(x) < 0 \quad \text{if} \quad x < c,$$

and

$$f'(x) > 0 \quad \text{if} \quad x > c,$$

then $f(x)$ is decreasing in $\{x : x < c\}$ and increasing in $\{x : x > c\}$. Clearly, $f(c)$ is the *overall minimum* of $f(x)$.



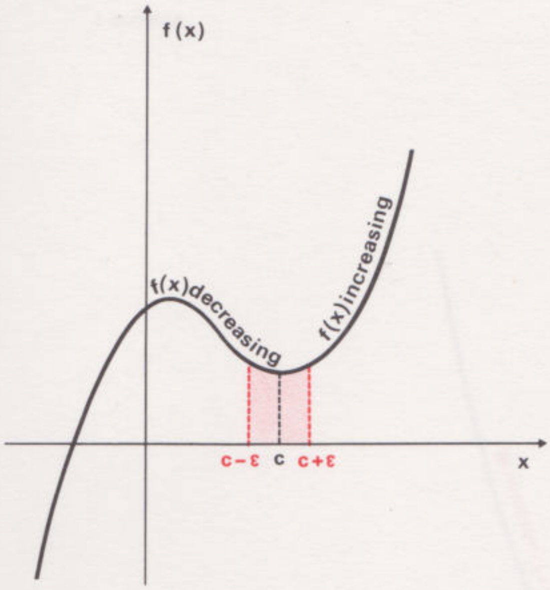
If there is a point $c \in R$ and a positive number ε such that

$$f'(x) < 0 \quad \text{if} \quad c - \varepsilon < x < c,$$

and

$$f'(x) > 0 \quad \text{if} \quad c < x < c + \varepsilon,$$

then, although we cannot say anything about overall maxima and minima, we can be sure that $f(c)$ is a *local minimum* of $f(x)$.



Exercise 1

Exercise 1
(3 minutes)

For each of the following functions determine the subset of the domain in which the images of the function are increasing and the subset in which they are decreasing.

(i) $f: x \longmapsto x^2$ $\quad (x \in [-5, 5])$
(ii) $\phi: x \longmapsto \pi x^2 - \pi x + \ln 3$ $\quad (x \in R)$
(iii) $g: t \longmapsto \dfrac{t^3}{3} - t + 2$ $\quad (t \in [-2, 3])$ ■

Exercise 2

Exercise 2
(3 minutes)

If

$$g: t \longmapsto \frac{t^3}{3} - t + 2 \qquad (t \in [-2, 3]),$$

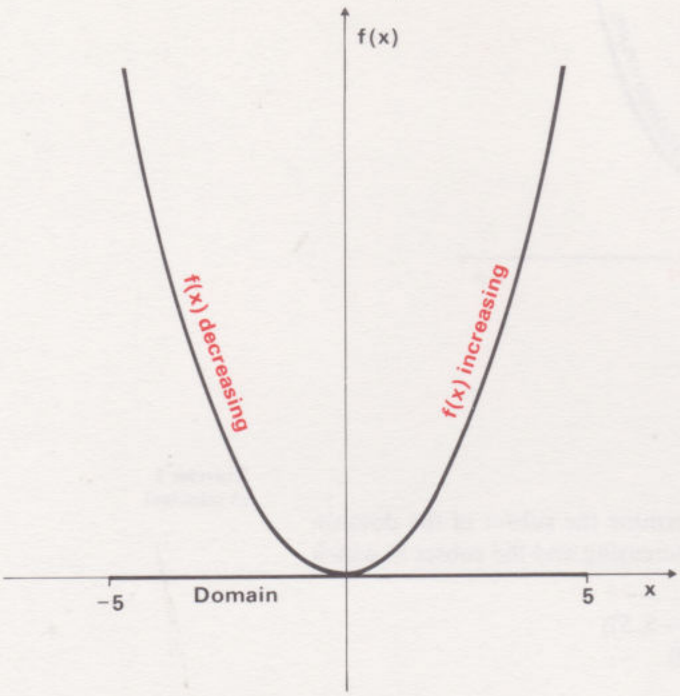
find a positive value of ε for which

(i) $g(t)$ is increasing in $]1, 1 + \varepsilon]$;
(ii) $g(t)$ is decreasing in $[1 - \varepsilon, 1[$;
(iii) $g(t) \geqslant g(1)$ for $t \in [1 - \varepsilon, 1 + \varepsilon]$. ■

A/515,33

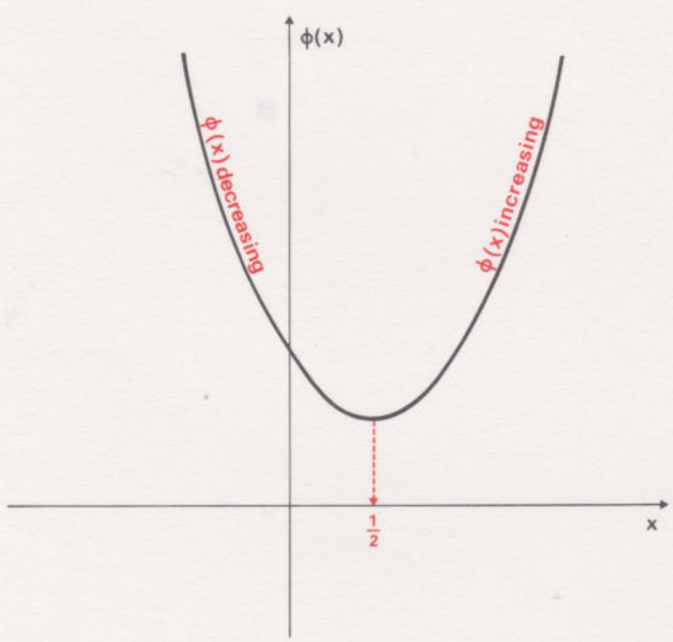
Solution 1 **Solution 1**

(i) $f'(x) = 2x$,
which is negative when $x < 0$
and positive when $x > 0$.



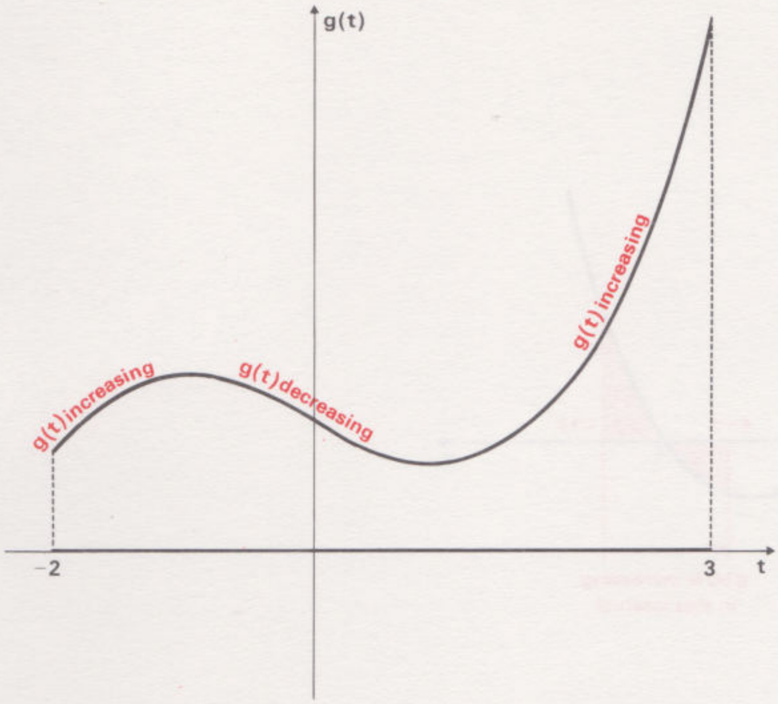
The required subsets are:
$[-5, 0[$ for f decreasing;
$]0, 5]$ for f increasing.

(ii) $\phi'(x) = 2\pi x - \pi$,
which is negative when $x < \frac{1}{2}$
and positive when $x > \frac{1}{2}$.



The required subsets are:
$\{x : x < \frac{1}{2}\}$ for ϕ decreasing;
$\{x : x > \frac{1}{2}\}$ for ϕ increasing.

(iii) $g'(t) = t^2 - 1$,
which is negative when $t^2 < 1$ (i.e. when $-1 < t < 1$)
and positive when $t^2 > 1$ (i.e. when $t < -1$ or $t > 1$).



The required subsets are:
$]-1, 1[$ for g decreasing;
$[-2, -1[\cup]1, 3]$ for g increasing. ■

Solution 2 **Solution 2**

For (i), any positive ε less than or equal to 2 will do.

For (ii), any positive ε less than 2 will do.

For (iii), any positive ε less than or equal to 2 will do. ■

Method Two

Main Text

You may be familiar with another technique which involves the second derived function of the function under consideration. This technique does, however, require a certain amount of care, not so much in its application, which is often very straightforward, but in the conclusions which you draw. Unfortunately, most students entirely ignore the method which we have just discussed once they are introduced to this second technique.

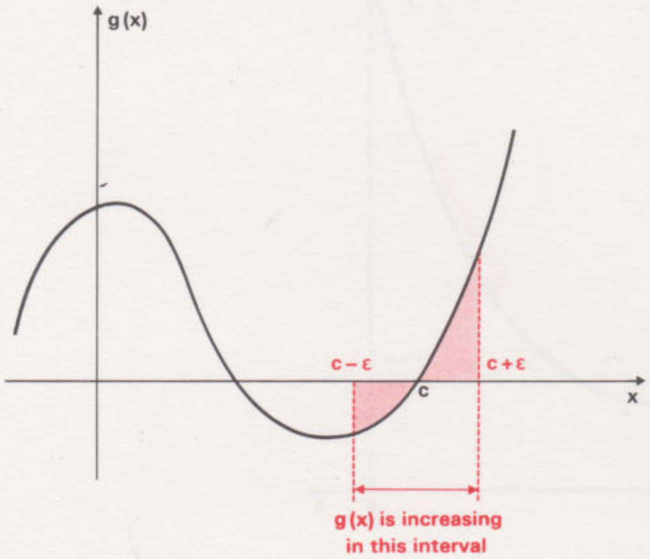
We know that the derived function, f', can be used to study the rate at which the images, $f(x)$, of a function f are changing. Similarly, f'' can be used to study the rate of change of $f'(x)$. In our previous method it was the fact that $f'(x)$ changed sign at a local maximum or minimum which was important, and it is a short step to the question: "Can we predict a change in the sign of $f'(x)$, if we are given f''?"

To find an answer to this question, suppose that we are given a real continuous differentiable function:

$$g : x \longmapsto g(x) \qquad (x \in R),$$

and that g' is also a real continuous function. (We shall say in a moment how g is related to f.) Suppose further that $g(c) = 0$ and $g'(c) > 0$ for some

$c \in R$; then, since g' is continuous, there must be an interval $[c - \varepsilon, c + \varepsilon]$ in which $g'(x) > 0$. It follows that $g(x)$ is increasing for $x \in [c - \varepsilon, c + \varepsilon]$. (As shown in the figure below, $g(x)$ may be increasing outside this interval as well, but that is immaterial to the argument.)



We see from the graph that

$$g(x) < g(c) \qquad \text{if} \quad c - \varepsilon < x < c,$$

and

$$g(x) > g(c) \qquad \text{if} \quad c < x < c + \varepsilon.$$

(Don't forget that one of our assumptions is that $g(c) = 0$.)

The useful piece of information which we are seeking follows from the above if we now assume that $g = f'$.

Notice that we require $g(c) = 0$, and this implies that

$$f'(c) = 0,$$

in other words, c is a stationary point of f. We also require that $g'(c) > 0$; since $g' = f''$, this implies that

$$f''(c) > 0.$$

With these conditions we were able to conclude that $g(x) < g(c)$ if $c - \varepsilon < x < c$, and $g(x) > g(c)$ if $c < x < c + \varepsilon$. Translated to give a result for f, this becomes

$$f'(x) < 0, \qquad \text{if } c - \varepsilon < x < c,$$

and

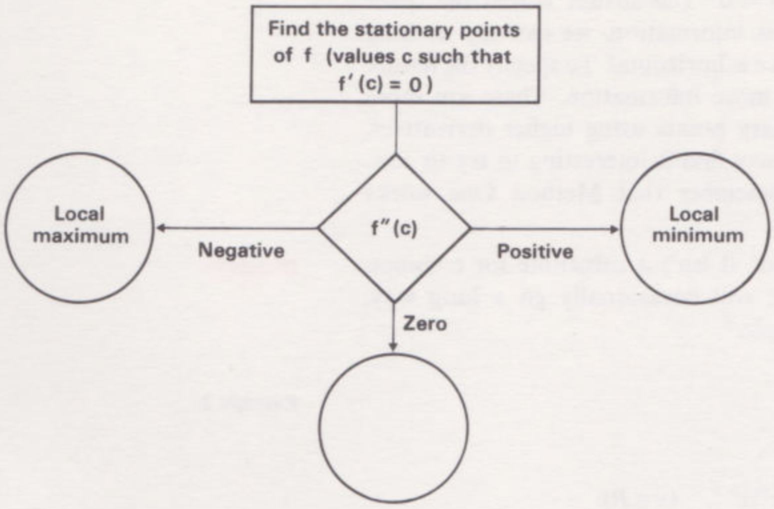
$$f'(x) > 0, \qquad \text{if } c < x < c + \varepsilon.$$

This means that there is a *local minimum* at c.

If originally we had taken $f''(c) < 0$, then our final conclusion would be that there is a *local maximum* at c.

(We assumed for convenience that the domain of f was R, but the same results are true for any function which has an interval as its domain.)

Classification of Stationary Points Using the Second Derivative



What should be entered in the blank circle?

A Few Words of Warning

Discussion

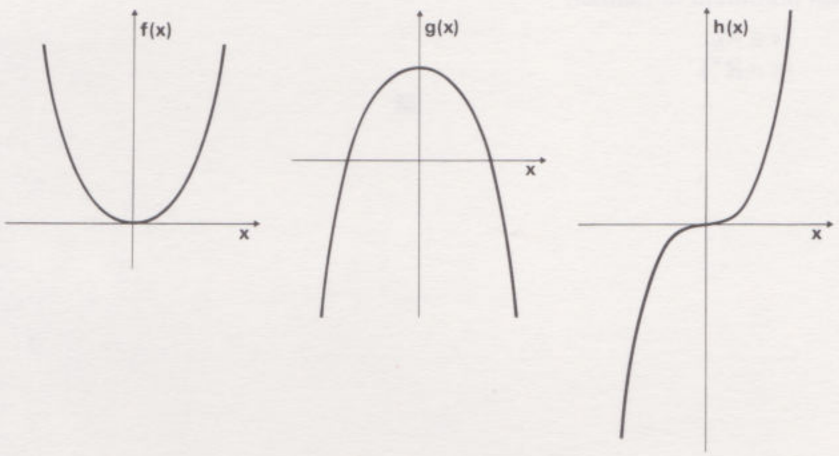
What can we say if $f'(c) = 0$ and $f''(c) = 0$? It is very tempting to say that f has neither a local maximum nor a local minimum at c, but this is wrong. The following examples should make the point clear.

Example 1

Example 1

Consider the following three functions, each with domain R:

$f: x \longmapsto x^4$ $\quad$ $g: x \longmapsto 1 - x^4$ $\quad$ $h: x \longmapsto x^3$



$f(x) = x^4$	$g(x) = 1 - x^4$	$h(x) = x^3$
$f'(x) = 4x^3$	$g'(x) = -4x^3$	$h'(x) = 3x^2$
$f'(0) = 0$	$g'(0) = 0$	$h'(0) = 0$
$f''(x) = 12x^2$	$g''(x) = -12x^2$	$h''(x) = 6x$
$f''(0) = 0$	$g''(0) = 0$	$h''(0) = 0$
From the graph, f has a local minimum at 0.	From the graph, g has a local maximum at 0.	From the graph, h has neither a local minimum nor a local maximum at 0.

■

What can we enter in the blank space on our flow chart corresponding to the case when $f'(c) = 0$ and $f''(c) = 0$? The answer is that the space is best left blank, for, given only this information, we can say nothing except that the tangent to the graph at c is horizontal. To specify the nature of the stationary point, we require more information. There are more powerful tests for classifying stationary points using higher derivatives, but we leave these until later. (You may find it interesting to try to construct such a test for yourself.) Remember that Method One works even when $f''(c) = 0$.

Discussion * *

Although calculus is a wonderful tool, it isn't a substitute for common sense. A little concentrated thought will occasionally go a long way, as you can see in the following example.*

Example 2

Example 2

Find the overall minimum value of

$$g(x) = ((x^4 + 2) + x^2(3 - x^2))^2 \qquad (x \in R).$$

If your first thought is: "Differentiate, and to the devil with the subtleties", then we admire your single-mindedness, but not your common sense.

The following solution is much simpler. Simplifying, we get

$$\begin{aligned} g(x) &= (x^4 + 2 + 3x^2 - x^4)^2 \\ &= (2 + 3x^2)^2. \end{aligned}$$

Since $x^2 \geqslant 0$, $2 + 3x^2$ takes its least value, 2, when $x = 0$; hence the overall minimum value of $g(x)$ is 4. ■

Exercise 3
(3 minutes)

Exercise 3

Find the stationary points of the following functions, and classify each of them as a local maximum, a local minimum or neither.

(i) $f: x \longmapsto x^3 - 6x^2 + 9x + 6 \qquad (x \in R)$,
(ii) $h: x \longmapsto x \ln x \qquad (x \in R^+)$.

■

* See also Exercise 15.1.2.1, part (i).

15.1.5 Examples and Exercises

15.1.5

Further Examples

You may omit this sub-section at a first reading if you wish, since it is intended only to consolidate the ideas introduced in the first section of this text. On the other hand, you may feel that you need more practice, in which case you should work through a number of examples and exercises until you feel confident enough to continue to the second and third sections.

Be careful not to spend too much time on any particular example or exercise; if there is a point which you don't understand, it is often a good idea to read on a little further.

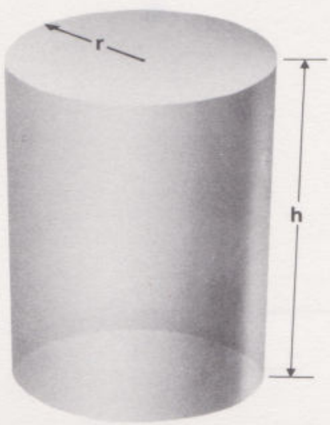
Example 1

Example 1

Of all the cylindrical tin cans (with lids) which enclose a volume of 1000 cm^3, which requires the least metal? ■

Solution of Example 1

Let S be the surface area of the can in cm^2; then, if the metal is of uniform thickness, the amount of metal required is proportional to S. Let h be the height, and r be the radius of an end-face measured in cm. (We have chosen the variables h and r, but the reason for this choice is not obvious. Choosing suitable notation is an essential prerequisite to mathematical solution, and we refer you to *Polya** (page 134) for a discussion of the matter.)



Since the volume enclosed is 1000 cm^3, we have

$$\pi r^2 h = 1000.$$

Equation (1)

Also, S is given by

$$S = 2\pi r^2 + 2\pi r h$$

Equation (2)

Our problem is to choose values of h and r satisfying Equation (1) which will minimize S.

Equation (2) could be used to define a function $(r, h) \longmapsto S$, with domain $R^+ \times R^+$, but in our problem r and h are not independent, and so we cannot choose (r, h) at random from $R^+ \times R^+$. By eliminating h, we can obtain a mapping

$$f: r \longmapsto S \qquad (r \in R^+),$$

* G. Polya, *How to Solve It*, Open University ed. (Doubleday Anchor Books 1970). This book is the set book for the Mathematics Foundation Course; it is referred to in the text as *Polya*.

(*continued on page 26*)

Solution 15.1.4.3

Solution 15.1.4.3

(i)
$$f'(x) = 3x^2 - 12x + 9$$
$$= 3(x-3)(x-1).$$

Thus $f'(x) = 0$ when $x = 1$ and when $x = 3$.

$$f''(x) = 6x - 12$$

and

$$f''(1) = -6,$$

which is less than 0, giving us a local maximum at $x = 1$.

$$f''(3) = 6,$$

which is greater than 0, giving us a local minimum at $x = 3$.

(ii)
$$h'(x) = x \times \frac{1}{x} + \ln x$$
$$= 1 + \ln x$$

Thus $h'(x) = 0$ when $\ln x = -1$, that is, when $x = \frac{1}{e}$.

$$h''(x) = \frac{1}{x},$$

which is greater than 0 when $x = \frac{1}{e}$, giving us a local minimum at this point; it is, in fact, an overall minimum. ■

(*continued from page 25*)

and then use the methods which we have developed to find the value of r which minimizes $S = f(r)$. Once r is determined, the corresponding value of h follows from Equation (1), and hence we know the dimensions of the required can.

Solving Equation (1) for h, we obtain $h = \frac{1000}{\pi r^2}$; then substituting for h in Equation (2) gives

$$S = 2\pi r^2 + 2\pi r\left(\frac{1000}{\pi r^2}\right)$$
$$= 2\pi\left(r^2 + \frac{1000}{\pi r}\right).$$

This equation defines a function:

$$f: r \longmapsto 2\pi\left(r^2 + \frac{1000}{\pi r}\right) \qquad (r \in R^+).$$

(Notice that we cannot have $r = 0$, for both practical and mathematical reasons.)

Differentiating f, we get

$$f'(r) = 2\pi\left(2r - \frac{1000}{\pi r^2}\right)$$
$$= 2\left(\frac{2\pi r^3 - 1000}{r^2}\right),$$

so $f'(r) = 0$ when $2\pi r^3 = 1000$.

The only stationary point of f is at

$$r = \frac{10}{\sqrt[3]{2\pi}}$$

Since

$$f'(r) < 0 \text{ when } 0 < r < \frac{10}{\sqrt[3]{2\pi}}$$

and

$$f'(r) > 0 \text{ when } \qquad r > \frac{10}{\sqrt[3]{2\pi}}$$

it follows from Method One (page 18)* that $f(r)$ takes its overall minimum value when

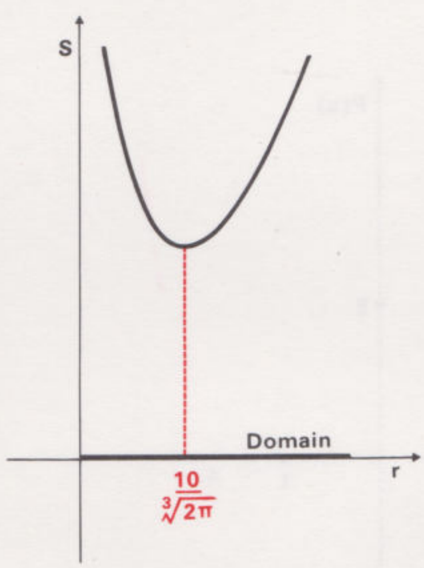
$$r = \frac{10}{\sqrt[3]{2\pi}}.$$

Substituting this value for r into Equation (1), we find that

$$h = \frac{20}{\sqrt[3]{2\pi}},$$

so that the tin which requires least metal has $h = 2r$.

The graph of f looks like this:



■

* You may like to show this by calculating $f''(r)$ and then substituting $r = \frac{10}{\sqrt[3]{2\pi}}$, if you feel that Method Two is easier.

Example 2

Example 2

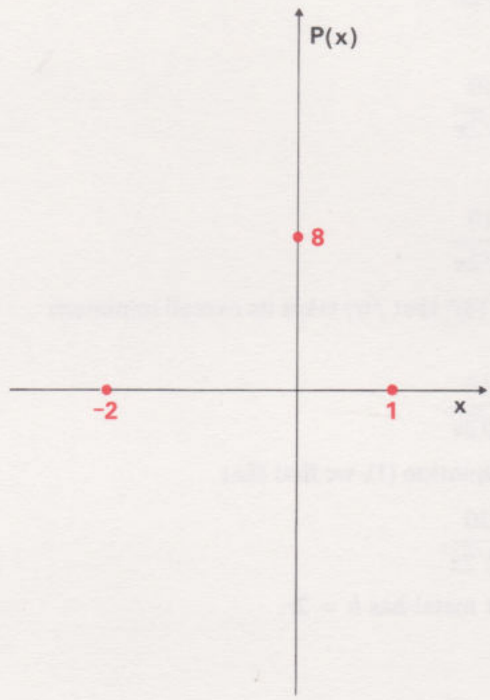
Sketch the graph of the function

$$P : x \longmapsto (x-1)^2(x+2)^3 \qquad (x \in R).$$

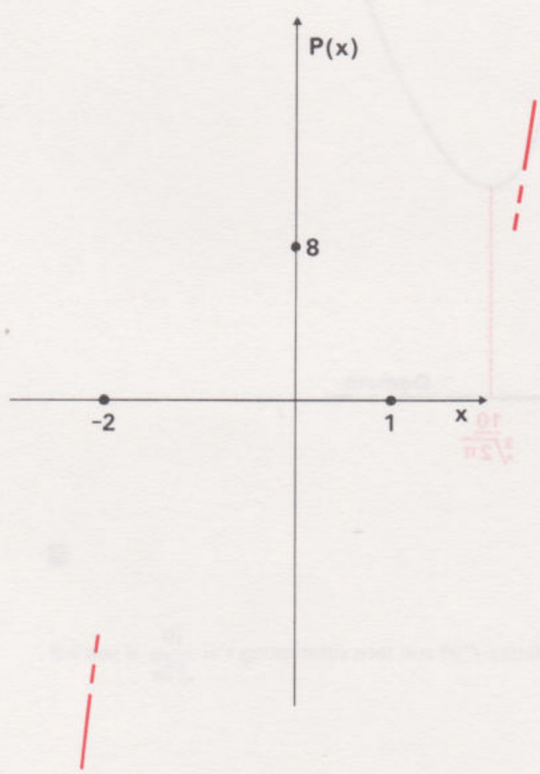
■

Solution of Example 2

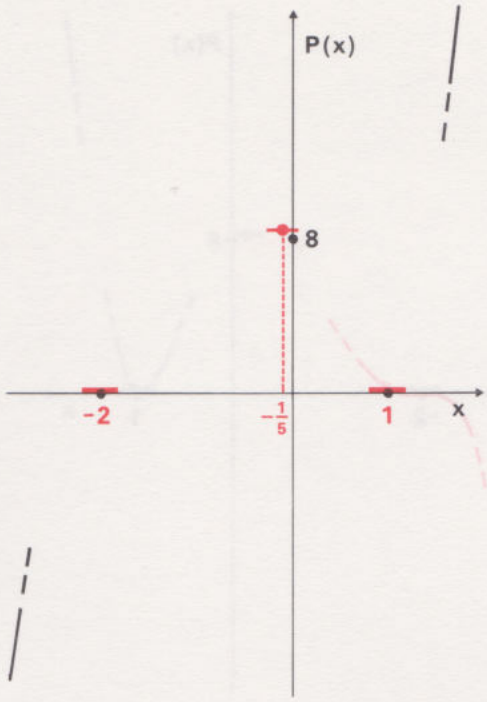
(i) Obviously $P(x) = 0$ if $x = 1$ or $x = -2$, and $P(x) = 8$ if $x = 0$. So these points are easily plotted on the graph.



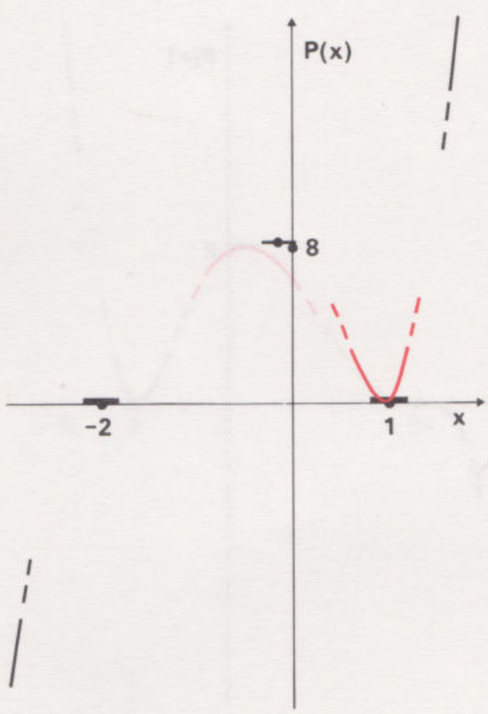
(ii) When x is a very large positive number, $P(x)$ is very large and positive. When x is a very large negative number, $P(x)$ is very large and negative.



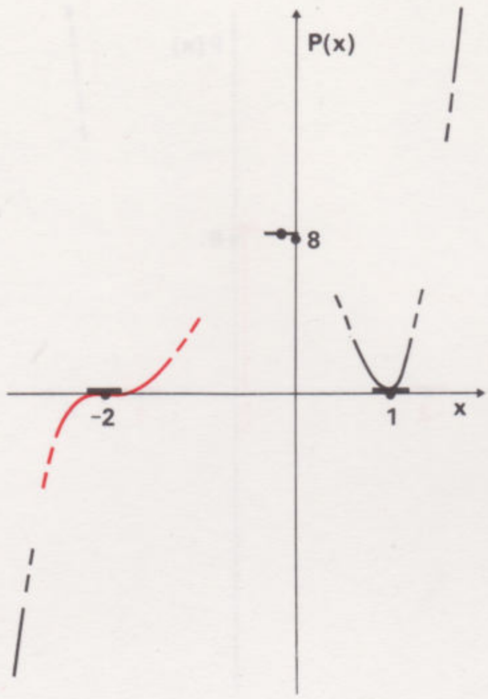
(iii) Differentiating P, we find (after a little manipulation which we leave as an exercise) that $P'(x) = (x - 1)(x + 2)^2(5x + 1)$, so that there are stationary points of P at 1, -2 and $-\frac{1}{5}$. It seems worth while to plot the point corresponding to $x = -\frac{1}{5}$.



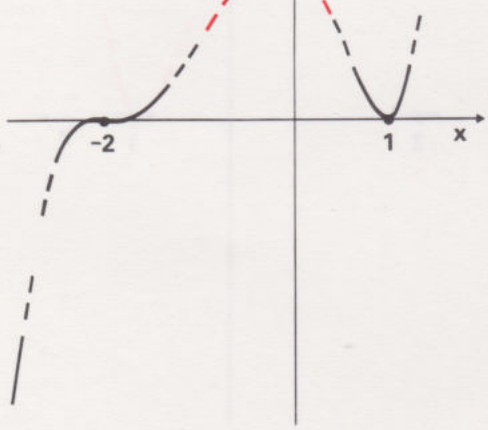
(iv) When x is any number slightly less than 1, the sign of $P'(x)$ is $(-)(+)(+) = (-)$, writing down only the signs of the various factors. When x is any number slightly greater than 1, the sign of $P'(x)$ is $(+)(+)(+) = (+)$. $P'(x)$ changes sign at $x = 1$, and therefore P has a local minimum at 1.



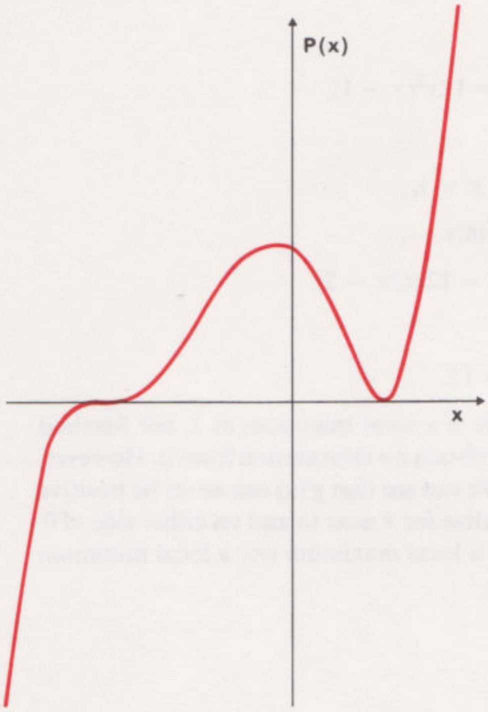
(v) When x is slightly less than -2, the sign of $P'(x)$ is $(-)(+)(-) = (+)$. When x is slightly greater than -2, the sign of $P'(x)$ is $(-)(+)(-) = (+)$. Hence P has neither a local maximum nor a local minimum at -2.



(vi) When x is slightly less than $-\frac{1}{5}$, the sign of $P'(x)$ is $(-)(+)(-) = (+)$. When x is slightly greater than $-\frac{1}{5}$, the sign of $P'(x)$ is $(-)(+)(+) = (-)$. Hence P has a local maximum at $-\frac{1}{5}$.



We can now be fairly confident that we can join the dotted line to complete the sketch.



■

Further Exercises

You will find exercises on this topic in every elementary book on calculus. In particular, there are a number of suitable exercises in Apostol, *Calculus Vol. 1* (see Bibliography), if you feel that you need more practice.

Exercise 1

Exercise 1
(3 minutes)

Find the stationary points of the following functions and classify each of them as a local maximum, a local minimum or neither.

(i) $g: x \longmapsto 3x^4 - 4x^3 \qquad (x \in R)$
(ii) $S: x \longmapsto \sin x \qquad (x \in [-\pi, \pi])$.

■

Solution 1

Solution 1

(i) $$g'(x) = 12(x^3 - x^2) = 12x^2(x - 1),$$

and

$$g'(x) = 0 \text{ if } x = 0 \text{ or } x = 1,$$

so these are the stationary points.

$$g''(x) = 12(3x^2 - 2x) = 12x(3x - 2),$$

so

$$g''(0) = 0 \text{ and } g''(1) = 12.$$

Immediately we can see that there is a local minimum at 1, but Method Two breaks down at 0 and we can obtain no information from it. However, our Method One will still work. We can see that $g'(x)$ can never be positive if $x < 1$, and so it is certainly negative for x near to and on either side of 0. It follows that g can have neither a local maximum nor a local minimum at 0.

(ii) $$S'(x) = \cos x,$$

and

$$S'(x) = 0 \text{ if } x = \pm\frac{\pi}{2},$$

so these are the stationary points,

$$S''(x) = -\sin x,$$

so

$$S''\left(-\frac{\pi}{2}\right) = 1, S''\left(\frac{\pi}{2}\right) = -1.$$

Hence we can deduce that S has a local maximum at $\frac{\pi}{2}$ and a local minimum at $-\frac{\pi}{2}$. ■

15.2 GEOMETRIC IDEAS

15.2

15.2.0 Introduction

15.2.0

Introduction

We shall continue to investigate the problem of optimization, but from now on we shall concentrate on functions of two (real) variables; that is, functions of the form:

$$F:(x, y) \longmapsto z \qquad ((x, y) \in R \times R),$$

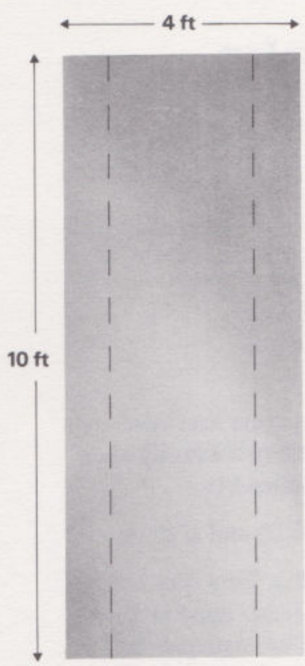
where $z \in R$. Before doing this we need to know a little three-dimensional co-ordinate geometry, because it is often helpful to represent such functions by surfaces. We really don't have enough time to do justice to geometry here; in fact a dedicated geometer would probably say that we were hardly doing geometry anyway. Our purpose in this section is to enable you to visualize the functions, and to describe the corresponding techniques in a pictorial and intuitive fashion. In section 15.3 we shall apply these geometric notions to our problem of optimization.

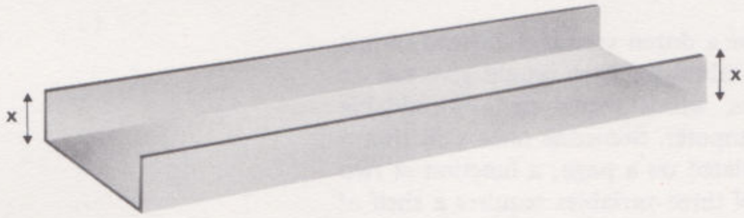
We begin by comparing a problem of optimizing a function of one variable with that of optimizing a function of two variables, just to give you a concrete example to cling to if you feel that things are getting too abstract. Then we shall see how various functions can be represented by surfaces. For our purposes, the most important of these surfaces is the plane, and the fact that we are able to use planes to investigate more complicated surfaces is crucial to the discussion which follows.

Example 1

Example 1

Suppose that an engineer is designing an aqueduct, and he has to use rectangular sheets of metal 4 ft wide by 10 ft long.





He intends to bend two edges at right-angles, as in the diagram, and his problem is to choose the value of x (in feet) which will allow the maximum amount of water to travel along the aqueduct.

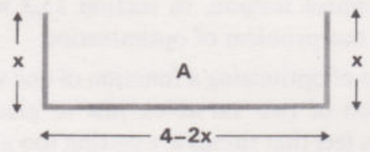
We shall try to find the value of the depth x which will give the greatest cross-sectional area, A (measured in square feet), of the channel. (It isn't obvious that this is the best thing to do; after all, the shape of the channel might alter the speed of the water flowing along it.) This merely requires a simple application of the techniques which we have developed in section 15.1.

Our problem in mathematical terms is to find the overall maximum value of the function:

$$f: x \longmapsto A \qquad (x \in [0, 2]).$$

We have

$$f(x) = x(4 - 2x),$$



and therefore

$$f'(x) = 4(1 - x),$$

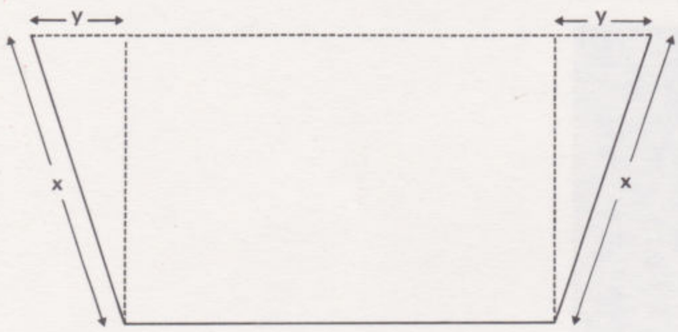
so that 1 is a stationary point.

$$f''(x) = -4$$

and therefore f has a local maximum at $x = 1$. It follows that the greatest possible value of A is

$$A = f(1) = 2.$$

The above solution is all very well, but wouldn't the engineer have been wiser to bend the sides at an angle other than a right angle?



In this case it shouldn't take you long to show that the cross-sectional area is $(4 - 2x + y)\sqrt{x^2 - y^2}$. This time our mathematical problem is to find the greatest value of $F(x, y)$, where the function F is defined by:

$$F: (x, y) \longmapsto (4 - 2x + y)\sqrt{x^2 - y^2} \qquad ((x, y) \in [0, 2] \times [0, 2] \text{ and } y \leqslant x).$$

The engineer now has two variables to juggle with, x and y. How can he solve a problem of this kind? He could use a common-sense approach, as we suggested initially for functions of one variable. For instance, he could take 10 values of y in all possible combinations with 10 values of x, and calculate the value of F corresponding to each pair; but this would mean 100 calculations, and if he took 50 values of each variable he would require $50^2 = 2\,500$ calculations.

In a practical situation we might have a dozen variables, instead of just two, in which case taking only ten values for each would give rise to $10^{12} = 1\,000\,000\,000\,000$ calculations, which would be a formidable task for even the fastest modern computer. Someone once said that a function of one variable can be tabulated on a page; a function of two variables needs a book; a function of three variables requires a shelf of

books; a function of four variables would take a whole library; and nobody in his right mind would try to tabulate a function of four or more variables.

Clearly we need something equivalent to the technique which we have just developed for functions of one variable.

We shall restrict the discussion in this unit to functions of two variables, although the basic results have equivalent forms for any number of variables. (A function of n real variables is a function which maps an element of the form $\underbrace{(x, y, \ldots, w)}_{n \text{ terms}}$ to a real number; that is, its domain is a subset of the Cartesian product, $\underbrace{R \times R \times \cdots \times R}_{n \text{ terms}}$ (which is usually denoted by R^n), and its codomain is R.) We shall discuss only functions of two variables, because such functions can be visualized in geometric terms.

15.2.1 Graphs of Functions of Two Variables

15.2.1

Introduction

We can represent many functions of one (real) variable by (pictorial) graphs, which enable us to use our intuition when examining the behaviour of the functions. In particular, when thinking of maxima and minima of such functions, we find the graphical approach very helpful. You should notice, however, that we try to discard the purely pictorial arguments in favour of symbolic reasoning, as soon as we feel that we are on the right track.

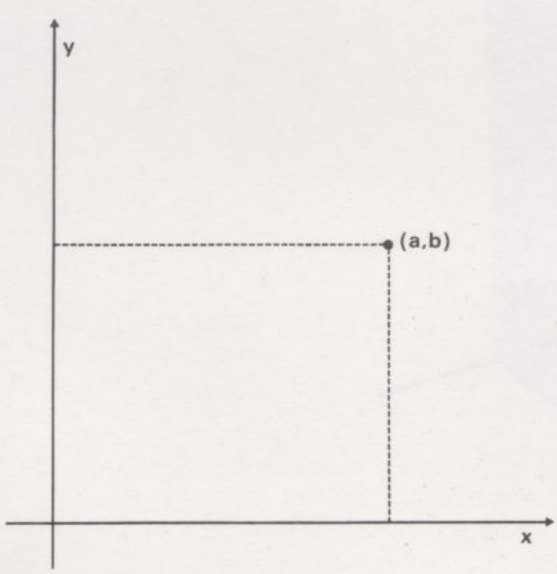
We shall base a number of arguments on pictures because we think that they are easier to understand this way: in later mathematics courses we shall need to examine more closely the difficulties which can arise.

Our first thought is to find a diagram which represents a function of two (real) variables, rather as a graph represents a function of one (real) variable. In this sort of diagram, we shall find that a function can often be represented by a surface. All our functions will be assumed to be "well-behaved"; in other words, the surfaces representing them have no spikes, gaps, or similar oddities.

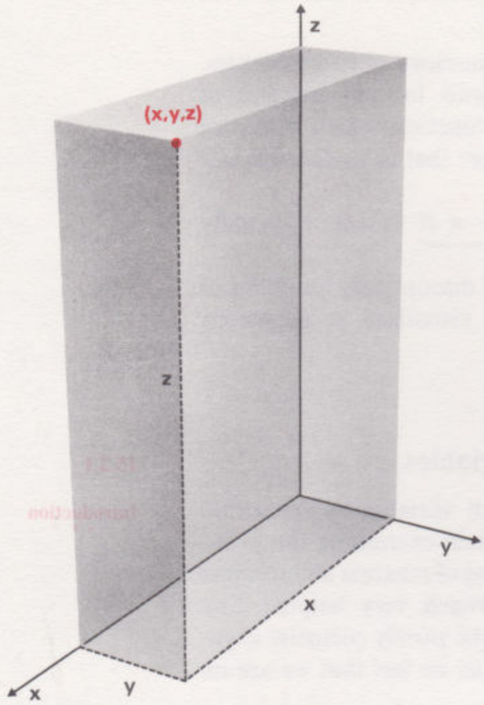
Cartesian Co-ordinates

Main Text

In the Cartesian plane each ordered pair of real numbers (a, b) corresponds to a unique point.



Likewise in the Cartesian space of three dimensions, each ordered triple (x, y, z) of real numbers corresponds to a unique point of the space.

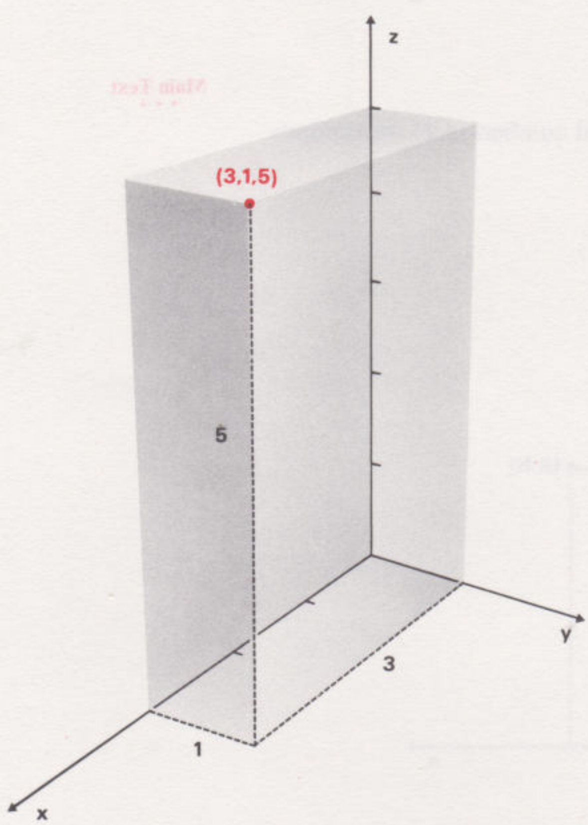


For example, we reach the point (3, 1, 5) if we start at the origin and proceed

3 units along the x-axis,
1 unit parallel to the y-axis,

and

5 units parallel to the z-axis.



Representation of Functions

Discussion

For functions of one variable, we know that a function gives rise to a graph (in the sense of a list), and this can be illustrated as a graph (in the sense of a picture). We now try to do something similar for a function of two variables. Let us look at a particular example of a function of two variables, and see how it gives rise first to a list and then to a picture.

Example 1

Example 1

Consider the function

$$F:(x, y) \longmapsto \sqrt{x^2 + y^2} \qquad ((x, y) \in R \times R).$$

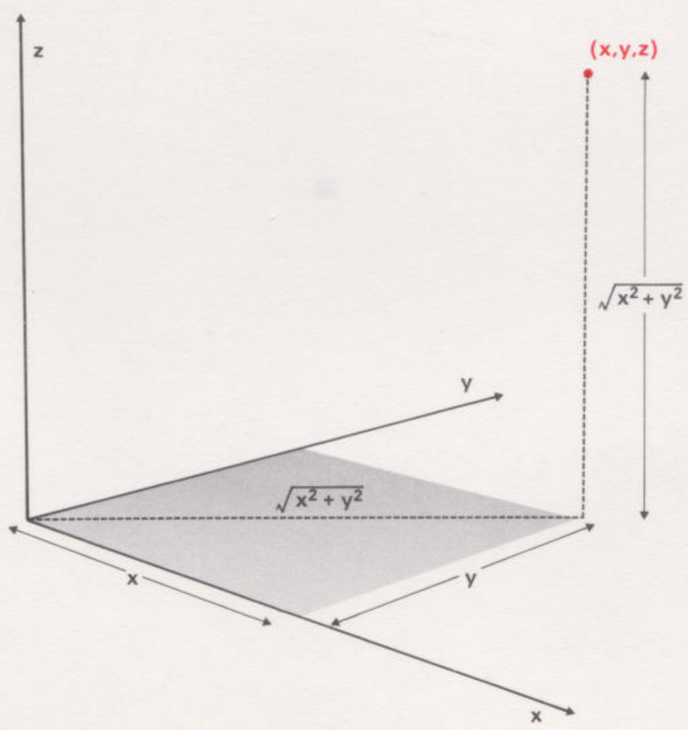
The ordered pair (3, 4) is mapped to $\sqrt{3^2 + 4^2} = 5$, and this corresponds to the ordered pair ((3, 4), 5). (Notice that the first element of this pair is also a pair.) Similarly, the pair (5, 12) maps to 13, and this corresponds to the pair ((5, 12), 13). If we put $F(x, y) = z$, then (x, y) maps to z, which gives rise to the pair $((x, y), z)$. In this way we can build up a table:

(x, y)	z
(3, 4)	5
(5, 12)	13
...	...

With the pair ((3, 4), 5) we can associate the point with co-ordinates (3, 4, 5); with the pair ((5, 12), 13) we can associate the point with co-ordinates (5, 12, 13); and so on. In this way, the function defines a set of ordered triples. Alternatively, we can think of the equation $z = F(x, y)$ as defining a *restriction* on the variables x, y and z. This restriction corresponds to a subset of $R \times R \times R$ (the set of all ordered triples of real numbers), namely the subset $\{(x, y, z): z = F(x, y)\}$.

The surface corresponding to this function F is particularly easy to visualize, for in this case

$$z = F(x, y) = \sqrt{x^2 + y^2}$$

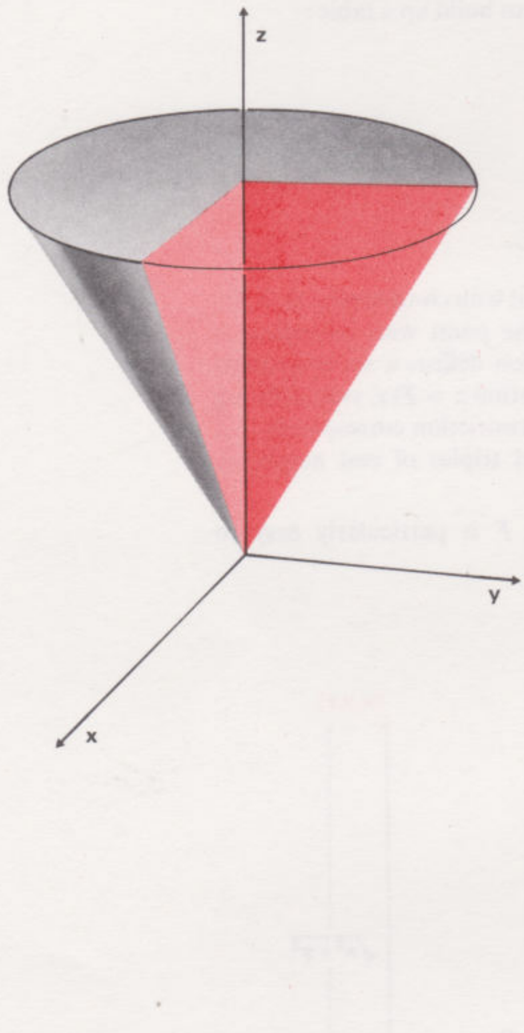


If we fix the value of z (corresponding to the vertical height in the diagram) and look at all the points at this height whose co-ordinates satisfy

$$z = \sqrt{x^2 + y^2},$$

then we find that they are all the same distance, $\sqrt{x^2 + y^2}$, from the z-axis; that is, they lie on a circle.

We can describe the surface in words by saying that we move from the origin in any horizontal direction, then vertically through the same distance to reach the surface. This surface is a cone with its vertex at the origin.



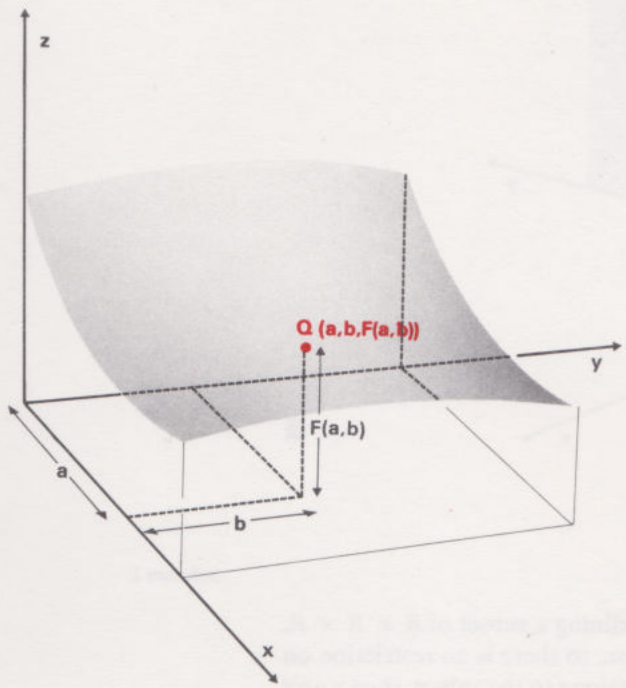
■

In general, let F be any function of two (real) variables:

$$F:(x, y) \longmapsto z \qquad ((x, y) \in R \times R).$$

Then to each ordered pair (a, b) we can associate the point Q with co-ordinates $(a, b, F(a, b))$.

In general, then, a function F of two variables defines a subset of $R \times R \times R$ which is often a surface. It is this surface which is the generalization of the pictorial graph of a function of one real variable.



Exercise 1

Exercise 1
(2 minutes)

Indicate on a diagram the sets of points with co-ordinates (x, y, z) satisfying:

(i) $x = 0$,
(ii) $y = 0$,
(iii) $x = y = 0$. ■

Exercise 2

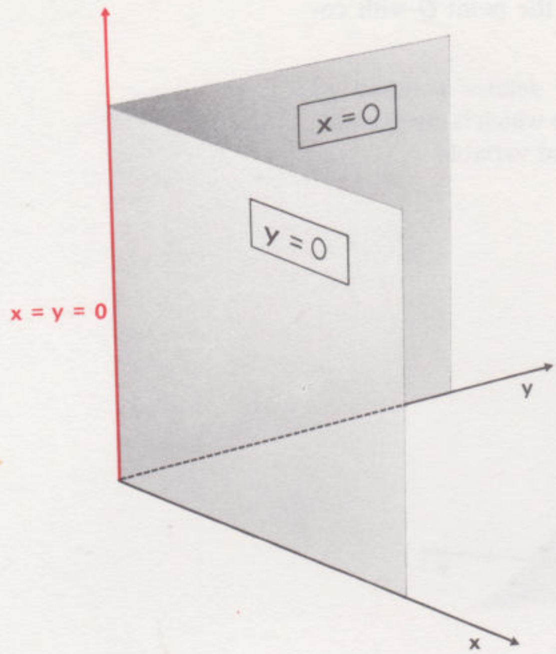
Exercise 2
(2 minutes)

Mark on a diagram the set of points with co-ordinates (x, y, z) in $R \times R \times R$ which corresponds to the condition:

$$2x - y = 0.$$ ■

Solution 1

Solution 1



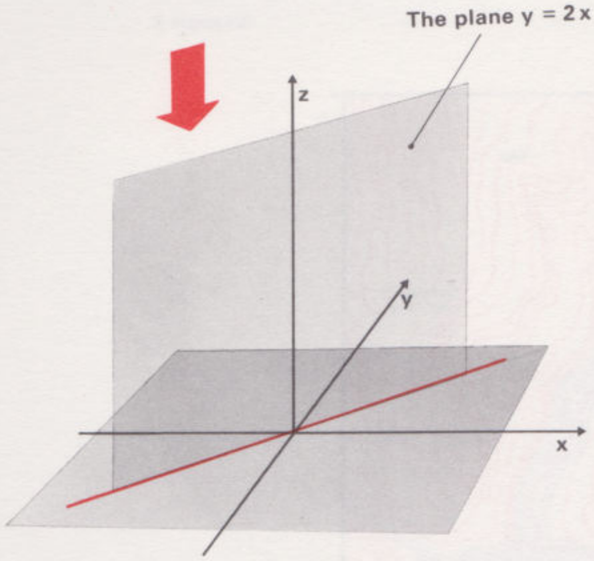
■

Solution 2

Solution 2

Considering the equation as a restriction defining a subset of $R \times R \times R$, we see that z does not appear in the equation, so there is no restriction on z. But x and y are restricted. If (x, y, z) is to belong to the subset, then x and y must satisfy the equation $2x - y = 0$. The set of triples $(x, y, 0)$ which satisfy this equation from a line in the xy-plane.

Corresponding to any point on this line, we can get other elements of the required subset of $R \times R \times R$ by choosing any value of z. All in all, we get a plane perpendicular to the xy-plane, which intersects the xy-plane in the line with equation $2x - y = 0$.



■

15.2.2 The Intersection of Surfaces

15.2.2

Discussion

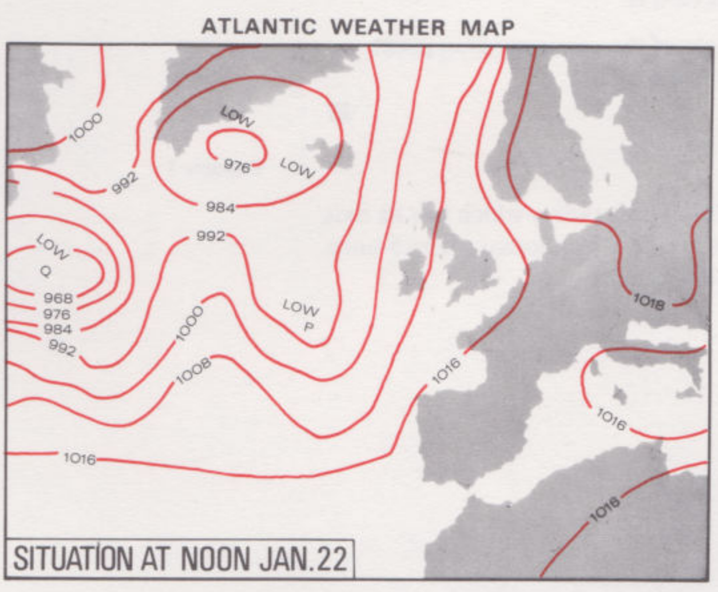
Before we find the equation which defines a general plane, we would like to give you some reasons for our interest in the subject. At first sight the following examples have nothing to do with planes, but a closer examination will reveal the connection.

Contour Lines

It is difficult to give an impression of a three-dimensional object on a two-dimensional piece of paper. One way of overcoming this difficulty, which is shown in the following examples, may be quite familiar to you, and we can develop it into a useful tool for examining functions of two variables.

Example 1 Barometric Pressure

Example 1

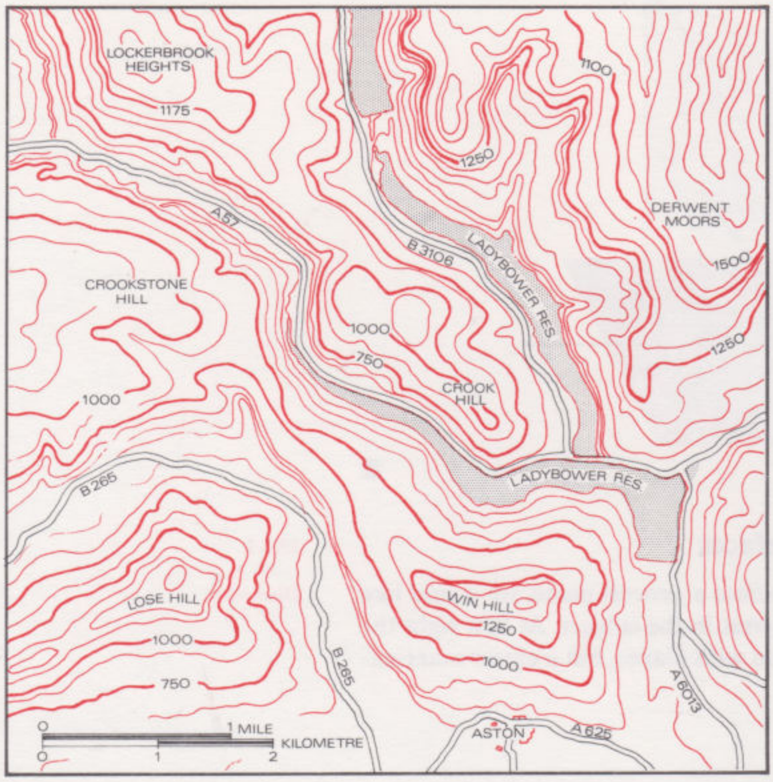


The red curves on the weather map join the points of equal barometric pressure. (The pressure varies with height, but the values shown refer to the pressure at sea-level.) The function illustrated in this case is

$$P: (\text{point on the map}) \longmapsto (\text{barometric pressure at the corresponding point on the earth's surface}).$$

■

Example 2 Ordnance Survey Maps **Example 2**



The cartographer has only a flat piece of paper, but he does his best to give an impression of the shape of the land surface by showing us the contour lines; in other words, he joins the points of equal height above sea level.

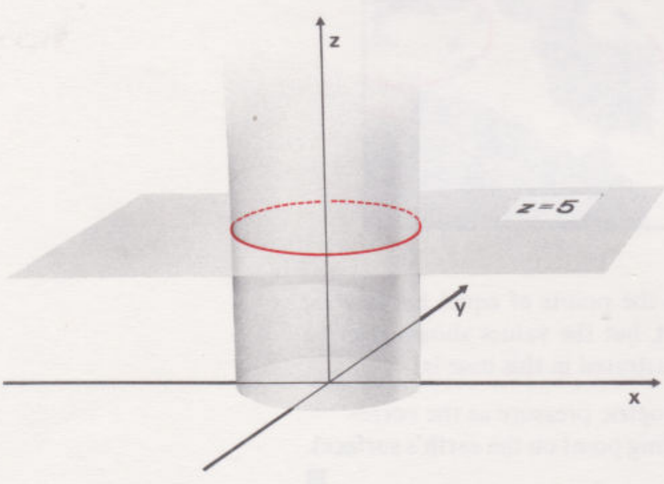
The function which is illustrated in this case is

$$h:(\text{point on the map}) \longmapsto (\text{height of the corresponding point above sea level})$$

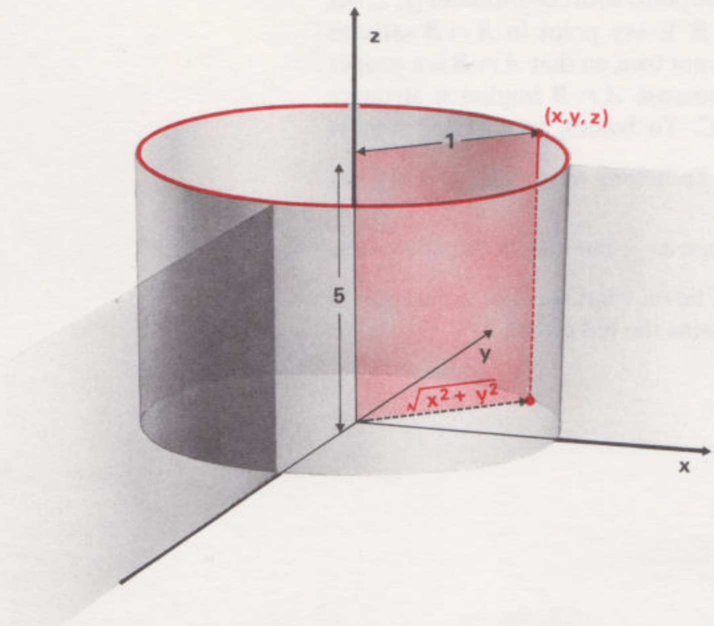
■

Example 3 **Example 3**

Suppose that we take a circular cylinder of unit radius which has its axis vertical (along the z-axis) and intersect it with a horizontal plane 5 units above the xy-plane, as shown in the following diagram.



The two surfaces (the cylinder and the plane) meet in a curve, which is in fact a horizontal circle 5 units above the xy-plane.



Any point on the cylinder is one unit from the z-axis and therefore $\sqrt{x^2 + y^2} = 1$. The equation of the cylinder is therefore $\sqrt{x^2 + y^2} = 1$, by which we mean that the set of all points with co-ordinates (x, y, z) in $R \times R \times R = R^3$, satisfying this equation, lie on the cylinder.

The equation of the plane is $z = 5$, and the two equations taken together:

$$\sqrt{x^2 + y^2} = 1$$

$$z = 5$$

determine the set of points lying on the red circle.

Another way of writing this is as follows: denote a point P with co-ordinates (x, y, z) by $P(x, y, z)$; then the cylinder is the set

$$A = \{P(x, y, z) : (x, y, z) \in R^3, \sqrt{x^2 + y^2} = 1\};$$

the plane is the set

$$B = \{P(x, y, z) : (x, y, z) \in R^3, z = 5\};$$

and the circle is $A \cap B$.

You may say that our two equations imply that $x^2 + y^2 = \frac{z}{5}$, so isn't this the equation of the circle? If so, compare the following two sets in $R \times R \times R$:

(i) $A \cap B$, the set of points with co-ordinates (x, y, z) which satisfy

$$\sqrt{x^2 + y^2} = 1$$

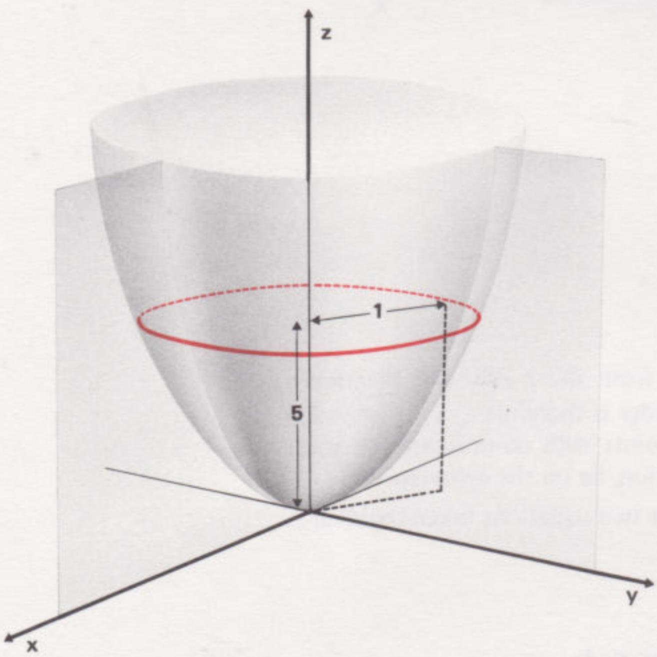
and

$$z = 5;$$

(ii) C, the set of points with co-ordinates (x, y, z) which satisfy

$$x^2 + y^2 = \frac{z}{5}.$$

Notice that every point in $A \cap B$ has $z = 5$, so that it must have co-ordinates of the form $(x, y, 5)$, whereas the point with co-ordinates $(2, 2, 40)$, for instance, lies in C but not in $A \cap B$. Every point in $A \cap B$ satisfies the conditions for C, but the converse is not true, so that $A \cap B$ is a proper subset of C. This is to be expected because $A \cap B$ implies a stronger restriction in $R \times R \times R$ than does C. To belong to C, (x, y, z) must be such that $x^2 + y^2$ is the same as $\frac{z}{5}$. To belong to $A \cap B$, (x, y, z) must be such that, not only is $x^2 + y^2$ the same as $\frac{z}{5}$, but also both expressions have the value 1. In fact the points in C lie on what is called a paraboloid of revolution, and this paraboloid contains the red circle.



■

A Generalization

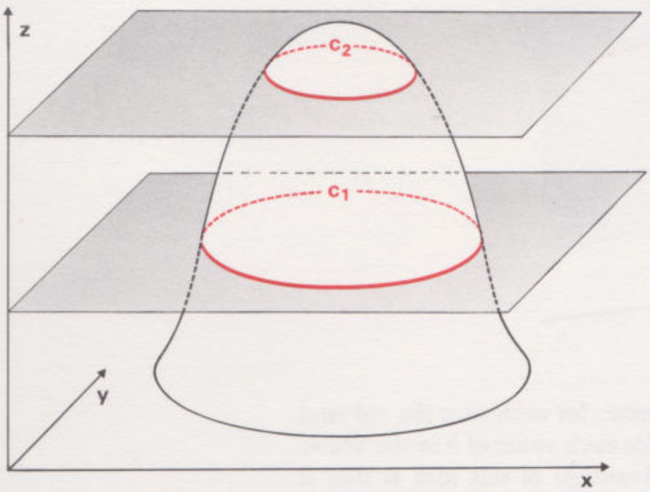
Discussion

Suppose that we are given an arbitrary function of two variables, F, with domain the xy-plane, and we intersect the surface

$$z = F(x, y)$$

with the horizontal plane

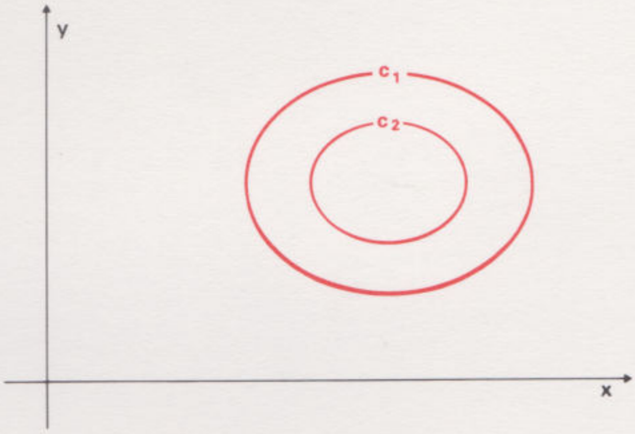
$$z = c$$



Various contour lines shown in red for various values of c.

The resulting curve is called the contour line corresponding to the height c. Taking various values of c will give a set of contour lines, which, when viewed from above (looking down the z-axis), could look like this:

Definition 1



This is simply a general version of the pressure and altitude diagrams which we used to introduce these ideas.

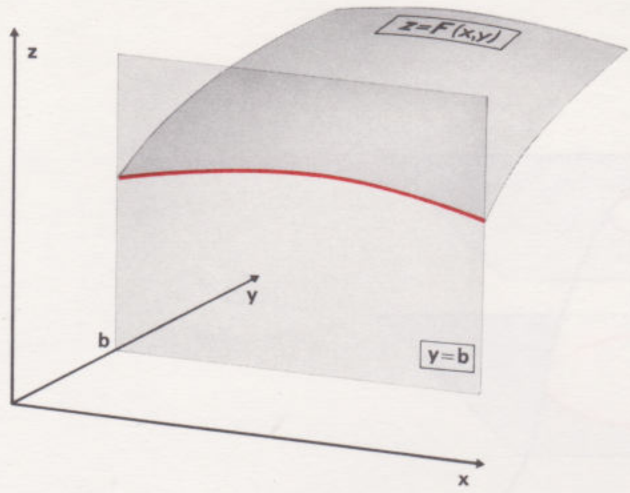
The previous examples have shown how planes parallel to the xy-plane (horizontal planes) can be used to describe surfaces, but we intend to use planes parallel to the z-axis (vertical planes) too.

Consider the intersection of our arbitrary surface defined by

$$z = F(x, y)$$

with the plane

$$y = b.$$



The effect is rather like slicing a Dutch cheese: for each slice the red rind of the cheese forms a different curve, and for each value of b in the above diagram we get a new red curve. The advantage of this idea is that it reduces a surface, which is difficult to draw, to a set of curves, each lying in a plane, which we *can* draw on a piece of paper. Mathematically speaking, we have reduced a function of two variables to a whole set of functions of one variable each corresponding to a particular value of b, and a particular red curve.

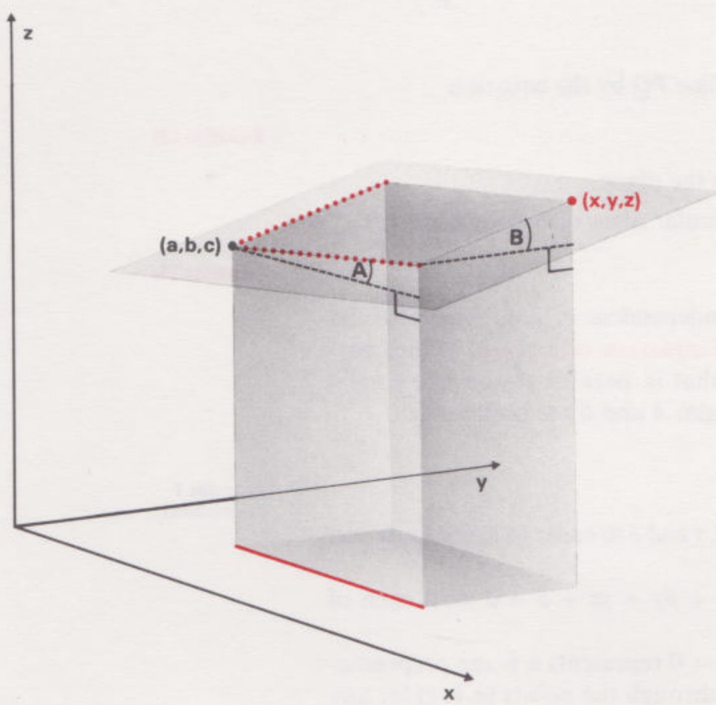
15.2.3 The General Equation of a Plane

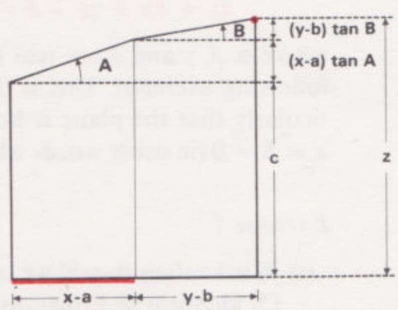
15.2.3

Discussion

Up to this point we have been finding the surfaces corresponding to given equations and functions, but now we want to put the problem in reverse. Can we find equations of given surfaces? What is the equation of a plane? If you have already seen the television programme, then you will know that this is an essential step on our way to solving optimization problems for functions of two variables.

Suppose that the plane passes through the point (a, b, c) and that it is inclined at an angle A in the x direction and an angle B in the y direction. You can see what this means from the cut-out diagram provided,* which, when set up with a sheet of paper placed on top to represent the plane, looks like this:





The value of z corresponding to an arbitrary choice of (x, y) is simply the result of adding the three terms on the right of the above diagram,

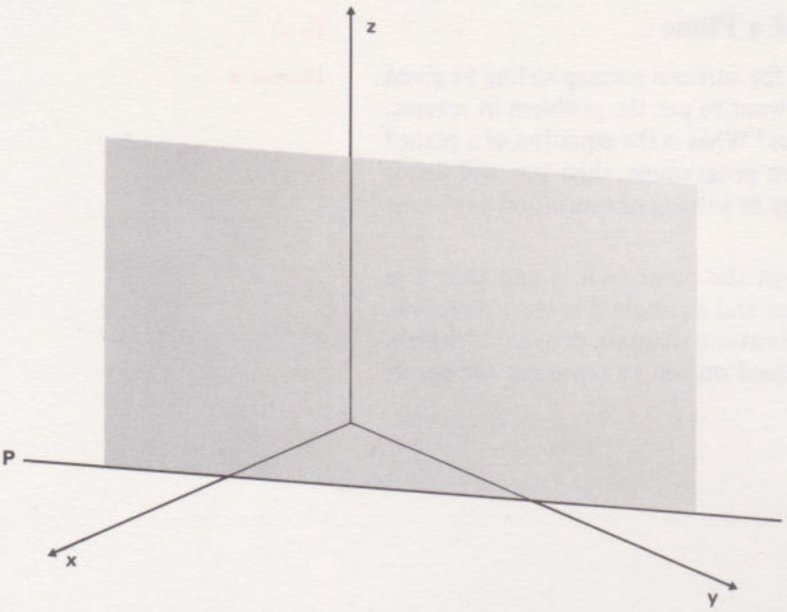
$$z = c + (x - a)\tan A + (y - b)\tan B,$$

Equation (1)

which is the required equation of the plane if neither A nor B is $\frac{\pi}{2}$.

If either A or B is $\frac{\pi}{2}$, then the plane is perpendicular to the xy-plane, and looking down the z-axis we see only the line PQ.

* The cut-out diagram is included inside the back cover of this text.



In the xy-plane we could represent the line PQ by the equation

$$\alpha x + \beta y + \delta = 0.$$ **Equation (2)**

In $R \times R \times R$ this equation represents the plane.

Equation (1) and Equation (2) are particular cases of the equation

$$\alpha x + \beta y + \gamma z + \delta = 0$$ **Equation (3)**

where α, β, γ and δ are real numbers independent of x, y, z (see also the following exercise). This is the general equation of a plane. Notice particularly that the plane is horizontal (that is, parallel to the xy-plane) if $\alpha = \beta = 0$ (in other words when the angles A and B are both zero).

Exercise 1

Exercise 1
(3 minutes)

(i) What values should we take for α, β, γ and δ in order to make Equation (3) identical to Equation (1)?
(ii) At what points does the plane $\alpha x + \beta y + \gamma z + \delta = 0$ meet each of the three co-ordinate axes?
(iii) The equation $\lambda(x - a) + \mu(y - b) = 0$ represents a plane perpendicular to the xy-plane which passes through the points (a, b, z) for any value of z. What effect does it have on the plane if we vary the values of λ and μ?

(λ and μ are the Greek letters called "lambda" and "mu" respectively.)

■

15.3 PARTIAL DERIVATIVES AND OPTIMIZATION OF FUNCTIONS OF TWO VARIABLES

15.3

15.3.0 Introduction

15.3.0

Introduction

We have derived the general equation of a plane, but if you have seen the television programme, you will know that we really need the equation of the *tangent plane* at a point on a given surface. We can then imagine this plane moving over the surface, and we hope that this notion will give us a technique for finding the maximum (or minimum) value of the corresponding function, just as a similar idea helped for functions of one variable. For the moment we need something like the derivative of a function of one variable, which was useful when discussing rate of change. The corresponding concept which we are going to examine is that of a *partial derivative*.

First let us give an intuitive idea of the concept of partial derivative. Imagine yourself standing at a crossroads on a hillside, the roads running East–West and North–South. Roughly speaking, the slopes of the East–West road and the North–South road are the partial derivatives of the function, represented by the hillside, at the point where the roads cross. If the crossroads happened to be at the top of a hill then each of the slopes would be zero. It is this intuitive idea that we want to make precise, and the following example will lead us in the right direction.

The geometric discussion in the example is intended to help you to understand the definitions which follow. Some people find three-dimensional figures hard to visualize, and if you don't like the geometry, you may be better off going directly to the definitions. You certainly should not spend a great deal of time trying to understand it if you find it difficult.

Example 1

Example 1

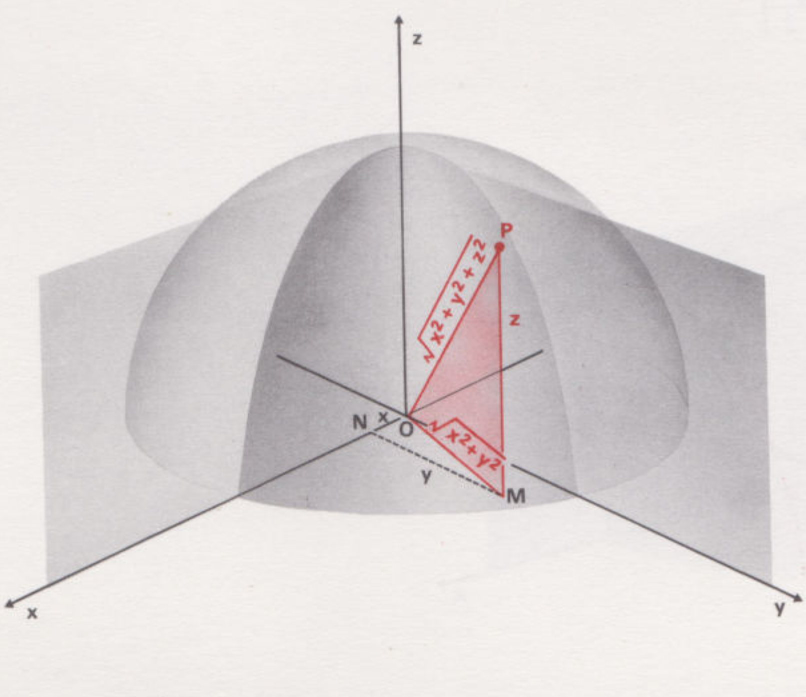
Consider the surface representing the function

$$F:(x, y) \longmapsto \sqrt{1-(x^2+y^2)} \qquad ((x, y) \in R \times R, x^2+y^2 \leqslant 1).$$

The domain of F is represented in the xy-plane by the points on and within the circle with radius 1, centred at the origin.

If we let $z = \sqrt{1-(x^2+y^2)}$, then it follows that $x^2+y^2+z^2 = 1$.

Equation (1)



(*continued on page 50*)

Solution 15.2.3.1

Solution 15.2.3.1

(i) $\alpha = \tan A$
$\beta = \tan B$
$\gamma = -1$
$\delta = c - a\tan A - b\tan B.$

(ii) If $\alpha \neq 0$, the plane meets the x-axis in the point with co-ordinates $\left(-\frac{\delta}{\alpha}, 0, 0\right)$.

If $\alpha = 0$, the equation of the plane is $\beta y + \gamma z + \delta = 0$:

(a) If $\delta \neq 0$, then y and z cannot be zero simultaneously, so the plane does not meet the x-axis.

(b) If $\delta = 0$, the equation of the plane is $\beta y + \gamma z = 0$. All points with co-ordinates $(x, 0, 0)$ lie on this plane, and so the plane contains the whole x-axis.

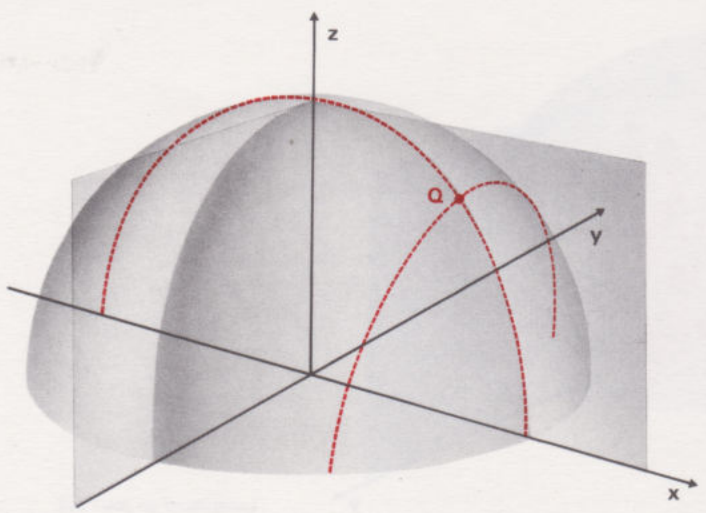
A similar argument can be used for the other co-ordinate axes.

(iii) Varying λ and μ alters the line in which the plane cuts the xy-plane. It will not change the fact that the plane is perpendicular to the xy-plane. ■

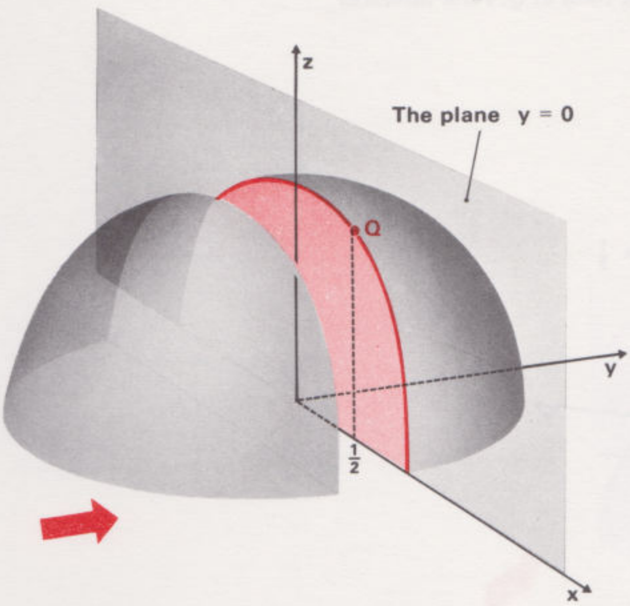
(*continued from page 49*)

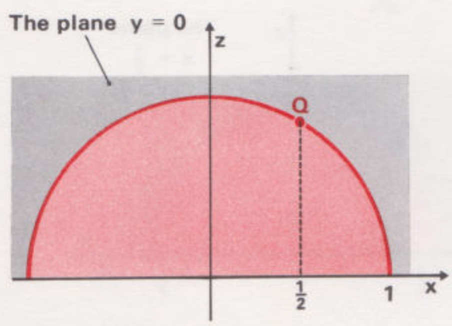
The distance of any point, $P(x, y, z)$, in $R \times R \times R$ from the origin is $\sqrt{x^2 + y^2 + z^2}$. This can be seen in the diagram, first by using Pythagoras's Theorem in the triangle ONM, and then in the triangle OMP. Since points on the surface satisfy the equation $x^2 + y^2 + z^2 = 1$, it follows that any point P lying on the surface must be at unit distance from the origin, and since z is always positive, Equation (1) represents a hemisphere.

In terms of our intuitive discussion, this hemisphere is the hillside. We are now going to choose a point Q on this hillside and assume that Q is our crossroads, with the roads through Q in the planes parallel to the x and y axes. Any point Q will do, and to illustrate the idea we choose Q to be the point with co-ordinates $\left(\frac{1}{2}, 0, \frac{\sqrt{3}}{2}\right)$.



The point Q with co-ordinates $\left(\frac{1}{2}, 0, \frac{\sqrt{3}}{2}\right)$ lies on the surface and on the plane $y = 0$. If we were to cut the hemisphere with the plane $y = 0$ through Q, and then look along the y-axis, we would see the semi-circle shown in red in the following diagram:





You can imagine the red curve to be the road through Q in the x-direction; later we shall find it very useful to be able to calculate the slope of such curves at an arbitrary point Q.

Next we find the slope of the red semi-circle at the point Q with co-ordinates $\left(\frac{1}{2}, 0, \frac{\sqrt{3}}{2}\right)$. The semi-circle is determined by the equations:

$$z = \sqrt{1 - (x^2 + y^2)}$$

$$y = 0$$

so that on the curve we have:

$$z = \sqrt{1 - x^2} \qquad (x \in [-1, +1]).$$

But a relationship like this defines a function, f_1 say:

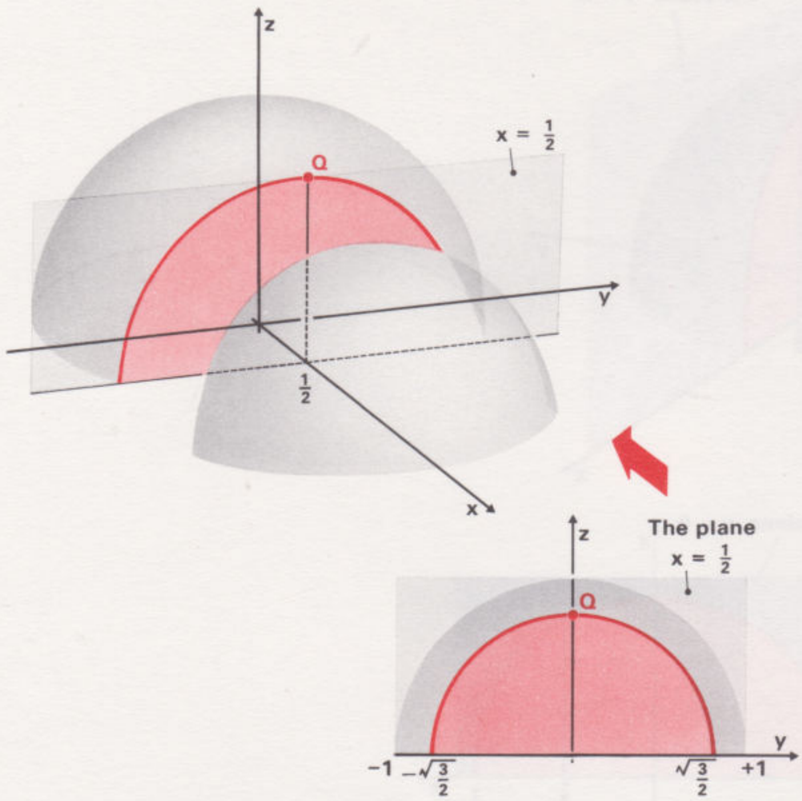
$$f_1 : x \longmapsto \sqrt{1 - x^2} \qquad (x \in [-1, +1]).$$

Since we are interested in the slope at Q, we calculate the derivative, $f'_1(x)$. We get

$$f'_1(x) = \frac{-x}{\sqrt{1 - x^2}}$$

which takes the value $-\dfrac{1}{\sqrt{3}}$ when $x = \dfrac{1}{2}$.

Suppose now that we want the slope of the other road at Q. First intersect the hemisphere with the plane $x = \frac{1}{2}$:



The red curve in the diagram is determined by the equations:

$$z = \sqrt{1 - (x^2 + y^2)}$$
$$x = \tfrac{1}{2}$$

so that on the curve we have:

$$z = \sqrt{1 - (\tfrac{1}{4} + y^2)}$$
$$= \sqrt{\tfrac{3}{4} - y^2} \qquad \left(y \in \left[-\frac{\sqrt{3}}{2}, +\frac{\sqrt{3}}{2}\right]\right).$$

This relationship defines a function, f_2 say:

$$f_2 : y \longmapsto \sqrt{\tfrac{3}{4} - y^2} \qquad \left(y \in \left[-\frac{\sqrt{3}}{2}, +\frac{\sqrt{3}}{2}\right]\right).$$

Again we calculate the derivative, $f'_2(y)$. We get

$$f'_2(y) = \frac{-y}{\sqrt{\frac{3}{4} - y^2}}$$

which takes the value 0 when $y = 0$ (the slope of the curve at Q seemed likely to be zero from the diagram, so this should be no surprise).

15.3.1 Definition of Partial Derivatives

15.3.1
Definitions
* * *

Having worked through the previous example, you may well feel that there must be a quicker way, and indeed there is. It is, however, the intuitive ideas of that example which point the way.

Suppose that we examine the function F from a slightly different point of view. We know that

$$F:(x, y) \longmapsto \sqrt{1-(x^2+y^2)} \qquad ((x, y) \in \{(x, y): x^2+y^2 \leqslant 1\}).$$

If we keep y constant, $y = b$ say, then we obtain a new function (of one variable):

$$f_1: x \longmapsto \sqrt{1-(x^2+b^2)} \quad (x \in [-\sqrt{1-b^2}, +\sqrt{1-b^2}]),$$

and

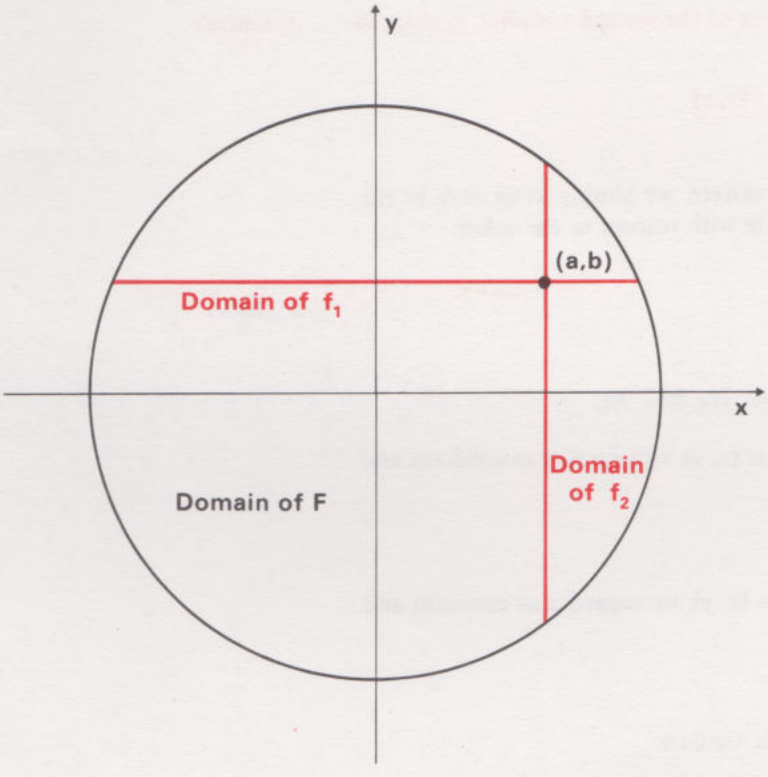
$$f_1'(x) = \frac{-x}{\sqrt{1-(x^2+b^2)}}.$$

Similarly, if we keep x constant, $x = a$ say, then we obtain a new function (of one variable):

$$f_2: y \longmapsto \sqrt{1-(a^2+y^2)} \quad (y \in [-\sqrt{1-a^2}, +\sqrt{1-a^2}]),$$

and

$$f_2'(y) = \frac{-y}{\sqrt{1-(a^2+y^2)}}.$$



The slopes of the roads through the point Q, $\left(\frac{1}{2}, 0, \frac{\sqrt{3}}{2}\right)$, running parallel to the x and y axes are given by the derivatives $f_1'(\frac{1}{2})$ and $f_2'(0)$ respectively (taking $a = \frac{1}{2}$ and $b = 0$).

The expression $f'_1(x)$ gives the slope of the surface (defined by the function F) in the direction of the x-axis at the point (x, b); that is, $f'_1(x)$ is the rate of change of F with respect to x, when y has the constant value b.

Similarly, $f'_2(y)$ gives the slope of the surface in the direction of the y-axis at the point (a, y); that is, $f'_2(y)$ is the rate of change of F with respect to y, when x has the constant value a.

We chose to consider the point Q with co-ordinates $\left(\frac{1}{2}, 0, \frac{\sqrt{3}}{2}\right)$, so we took $a = \frac{1}{2}$ and $b = 0$. We would like to know the corresponding slopes (rates of change) at *any* point on the surface defined by F; that is, we now wish to *vary* a and b. This means that we need to express the slopes in terms of functions of *two* variables. So we define two *new* functions, F'_1 and F'_2, by the equations:

$$F'_1(x, y) = \frac{-x}{\sqrt{1-(x^2+y^2)}} \qquad ((x, y) \in \{(x, y): x^2 + y^2 \leqslant 1\})$$

and

$$F'_2(x, y) = \frac{-y}{\sqrt{1-(x^2+y^2)}} \qquad ((x, y) \in \{(x, y): x^2 + y^2 \leqslant 1\}).$$

We are thus led to the following *definition of partial derivatives* of a function, F, of two variables x and y.

Definition 1

The partial derivative of F with respect to the first variable, x, at (x, y) is

$$F'_1(x, y) = \lim_{h \to 0} \frac{F(x+h, y) - F(x, y)}{h}$$

Definition 2

The partial derivative of F with respect to the second variable, y, at (x, y) is

$$F'_2(x, y) = \lim_{k \to 0} \frac{F(x, y+k) - F(x, y)}{k}$$

In order to find the two partial derivatives, we simply keep each of the variables fixed in turn and differentiate with respect to the other.

Example 1 **Example 1**

If

$$G: (x, y) \longmapsto 2xy + x^2 \qquad ((x, y) \in R \times R),$$

then, differentiating with respect to x at (x, y), we regard y as constant and get

$$G'_1(x, y) = 2y + 2x,$$

and differentiating with respect to y at (x, y), we regard x as constant and get

$$G'_2(x, y) = 2x.$$

(Working directly from the definitions, we have:

$$\begin{aligned} G'_1(x, y) &= \lim_{h \to 0} \left(\frac{2y(x+h) + (x+h)^2 - (2xy + x^2)}{h} \right) \\ &= \lim_{h \to 0} \left(\frac{2yh + 2xh + h^2}{h} \right) \\ &= 2y + 2x, \end{aligned}$$

and

$$G'_2(x, y) = \lim_{k \to 0} \left(\frac{2x(y + k) + x^2 - (2xy + x^2)}{k} \right)$$

$$= \lim_{k \to 0} \left(\frac{2xk}{k} \right)$$

$$= 2x \qquad \blacksquare$$

Exercise 1

Exercise 1
(2 minutes for each part)

Find the partial derivatives at (x, y) of the functions defined by the following equations; each function has domain $R \times R$.

(i) $F(x, y) = x^2 + y^2$
(ii) $G(x, y) = x \exp(xy)$
(iii) $H(x, y) = x \sin(x + y)$
(iv) $P(x, y) = x^4 + y^4 - 4x^2y^3$. $\blacksquare$

Alternative Notation

Notation

There are various notations for the partial derivatives; the most common is $\frac{\partial F}{\partial x}$ for what we write as $F'_1(x, y)$. This notation arose presumably because of the commonly used notation $\frac{df}{dx}$ for the derivative of a function, f, of one variable. If you use the $\frac{\partial F}{\partial x}$ notation, then you must be extremely careful later. For example, it is not generally true that $\frac{\partial F}{\partial x} \cdot \frac{\partial x}{\partial t}$ is simply $\frac{\partial F}{\partial t}$, as the notation would suggest.

The alternative notation F_x or F_y is also very common.

Solution 1

Solution 1

(i) $F_1'(x, y) = 2x$
$F_2'(x, y) = 2y$

(ii) $G_1'(x, y) = \exp(xy) + xy\exp(xy)$
$G_2'(x, y) = x^2\exp(xy)$

(iii) $H_1'(x, y) = \sin(x + y) + x\cos(x + y)$
$H_2'(x, y) = x\cos(x + y)$

(iv) $P_1'(x, y) = 4x^3 - 8xy^3$
$P_2'(x, y) = 4y^3 - 12x^2y^2$. ■

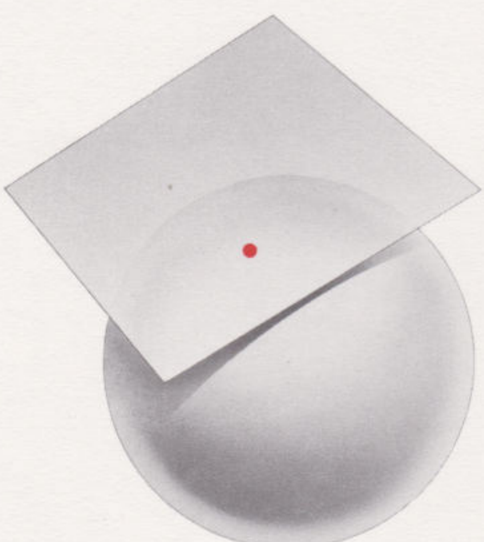
15.3.2 The Tangent Plane

15.3.2

Discussion

You probably have an intuitive notion of what we mean by the *tangent plane* at a particular point on a surface. It is, after all, the plane which sits comfortably on the surface at the point in question. Once again, we assume that our surfaces are smooth with no sharp projections. It would, for example, be difficult to decide where the tangent plane should be on the apex of a church steeple.

On the other hand, it is quite easy to imagine a tangent plane at a point on a smooth sphere.



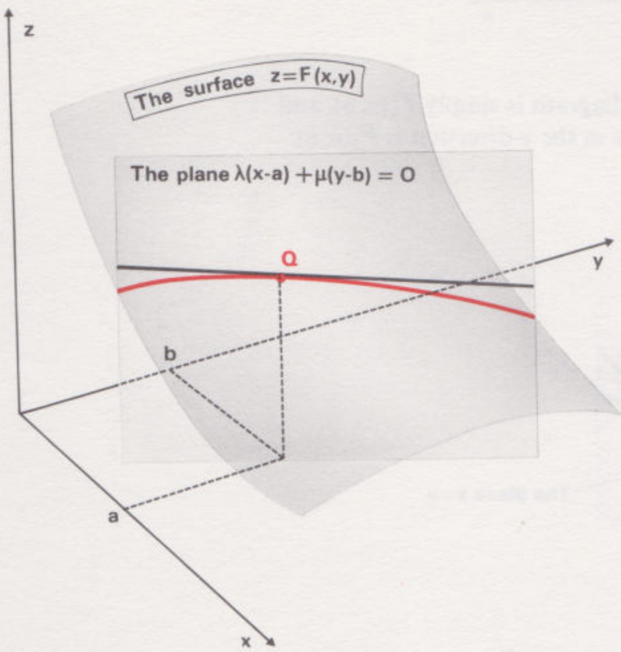
We shall now define the tangent plane at any point on a smooth surface. Suppose that we are given a surface defined by:

$$F:(x, y) \longmapsto F(x, y) \qquad ((x, y) \in R \times R),$$

and we wish to define the tangent plane at the point Q with co-ordinates $(a, b, F(a, b))$. We have seen in Exercise 15.2.3.1 (iii) that the equation

$$\lambda(x - a) + \mu(y - b) = 0$$

defines a plane which passes through Q and is perpendicular to the xy-plane.



The intersection of this plane with the surface will be a curve (shown in red on the diagram). This curve passes through Q and has a tangent line (shown by a heavy black line) at Q. If we vary the values of λ and μ, the plane will rotate about the vertical line through Q, and each pair of values of λ and μ will give us such a tangent line. If *all* these tangent lines at Q lie in a plane, then we call this plane the tangent plane at Q. **Definition 1**

The Equation of the Tangent Plane

Main Text

Our assumption that the surfaces we meet are smooth is intended to imply that there *is* a tangent plane to the surface $z = F(x, y)$ at Q, but how can we find its equation?

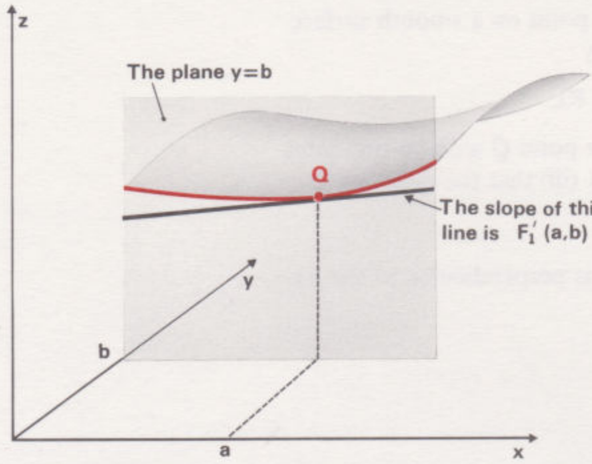
Suppose that we take the particular values $\mu = 1$, $\lambda = 0$, in the equation $\lambda(x - a) + \mu(y - b) = 0$.

Then we simply get the equation

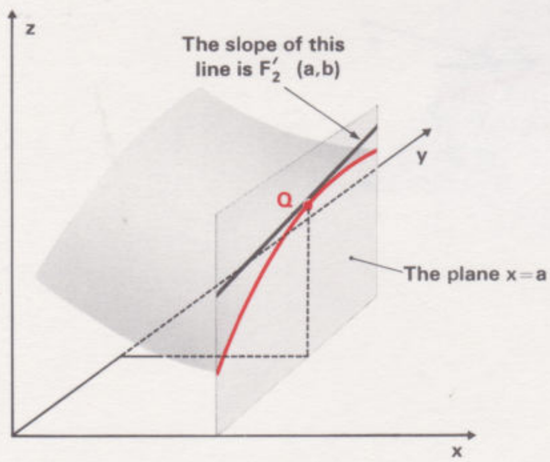
$$y = b,$$

and the slope of the corresponding curve of intersection at Q is $F_1'(a, b)$. In other words, $F_1'(a, b)$ is the slope of the tangent to this curve at Q.

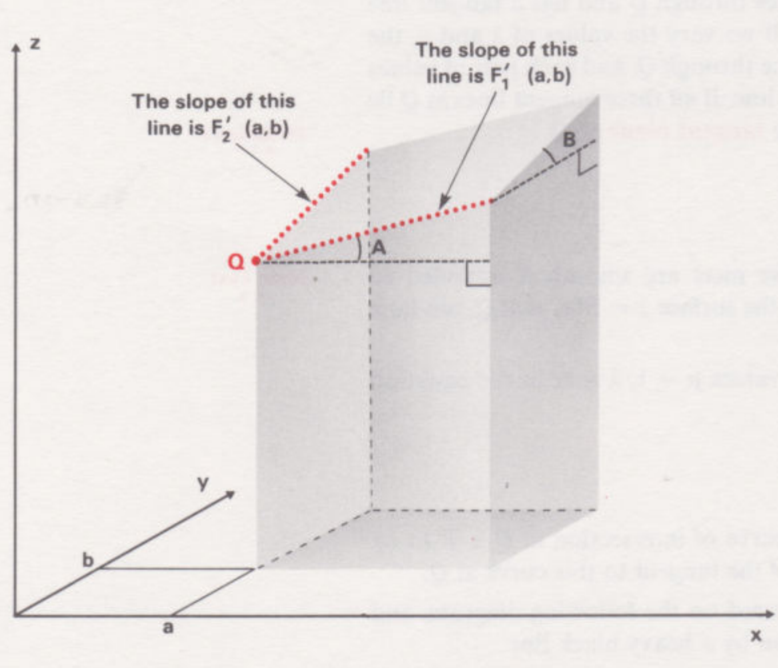
The curve of intersection is shown in red on the following diagram, and the tangent to this curve at Q is shown by a heavy black line.



The slope of the tangent line in the above diagram is simply $F'_1(a, b)$, and similarly the slope of the corresponding line in the y-direction is $F'_2(a, b)$.



We can now use our cut-out model of a plane again, this time to find the equation of the tangent plane at Q.



In the formula on page 47, we simply put $\tan A = F_1'(a, b)$, $\tan B = F_2'(a, b)$ and $c = F(a, b)$. We can see from the last diagram that any point on the tangent plane has co-ordinates (x, y, z) which satisfy the equation

$$z = F(a, b) + F_1'(a, b)(x - a) + F_2'(a, b)(y - b)$$

Equation of Tangent Plane

* * *

and this is the equation of the tangent plane to the surface at $(a, b, F(a, b))$

Exercise 1

Exercise 1
(2 minutes for each part)

For each of the following functions, find the equation of the tangent plane at the point on the surface corresponding to the pair (a, b) (each function has domain $R \times R$).

(i) $F:(x, y) \longmapsto x^2 + y^2$
(ii) $G:(x, y) \longmapsto x \exp(xy)$
(iii) $H:(x, y) \longmapsto x \sin(x + y)$
(iv) $P:(x, y) \longmapsto x^4 + y^4 - 4x^2y^3$.
(You will be able to use the results of Exercise 15.3.1.1.) ■

Exercise 2

Exercise 2
(3 minutes)

Find the equation of the tangent plane to the surface defined by

$$F:(x, y) \longmapsto (4 - 2x + y)\sqrt{x^2 - y^2}$$
$$((x, y) \in \{(x, y): 0 \leqslant y \leqslant x \leqslant 2\})$$

at the point $(a, b, F(a, b))$. (This is closely related to the aqueduct problem which we discussed in section 15.2.0.) ■

Solution 1 **Solution 1**

(i) $z = a^2 + b^2 + 2a(x - a) + 2b(y - b)$
(ii) $z = a\exp(ab) + (ab\exp(ab) + \exp(ab))(x - a) + a^2\exp(ab)(y - b)$
(iii) $z = a\sin(a + b) + (\sin(a + b) + a\cos(a + b))(x - a) + a\cos(a + b)(y - b)$
(iv) $z = a^4 + b^4 - 4a^2b^3 + (4a^3 - 8ab^3)(x - a) + (4b^3 - 12a^2b^2)(y - b)$. ■

Solution 2 **Solution 2**

First, we must find the partial derivatives; they are:

$$F_1'(x, y) = -2\sqrt{x^2 - y^2} + (4 - 2x + y)\frac{x}{\sqrt{x^2 - y^2}}$$

$$F_2'(x, y) = \sqrt{x^2 - y^2} - (4 - 2x + y)\frac{y}{\sqrt{x^2 - y^2}}$$

Thus the equation of the tangent plane is

$$z = (4 - 2a + b)\sqrt{a^2 - b^2} + \left(-2\sqrt{a^2 - b^2} + \frac{(4 - 2a + b)}{\sqrt{a^2 - b^2}}a\right)(x - a) + \left(\sqrt{a^2 - b^2} - \frac{(4 - 2a + b)}{\sqrt{a^2 - b^2}}b\right)(y - b)$$

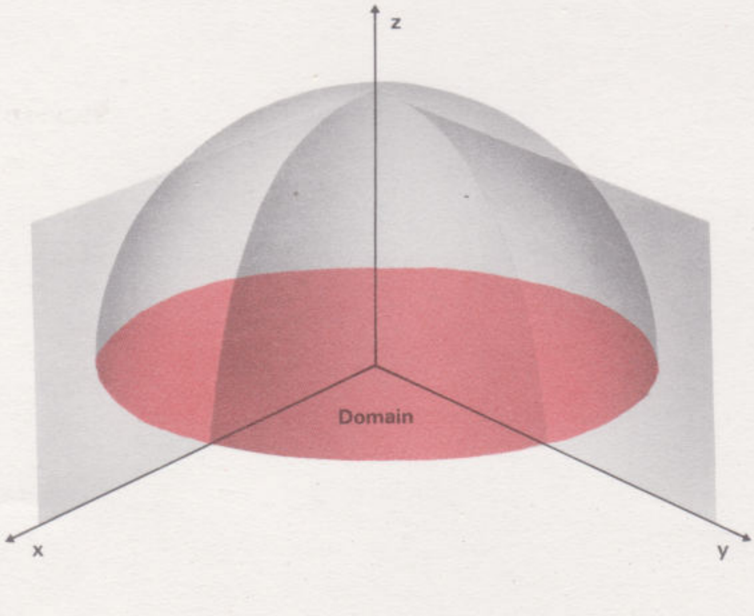
Unfortunately, this cannot be simplified; we included the exercise because of its tie-up with the aqueduct problem. ■

15.3.3 Optimizing Functions of Two Variables

15.3.3

Discussion

There is no doubt that finding the overall maximum (or minimum) of a function of two variables is, in general, harder than finding the overall maximum (or minimum) of a function of one variable. Some might say, "more than twice as hard". For a differentiable function f of one variable with domain $[a, b]$ we merely locate the stationary points, investigate their nature, and then examine the values of $f(a)$ and $f(b)$. The domain of a (real) function of two variables is a subset of $R \times R$, and instead of just two end-points we are now likely to have a curve as the boundary of our domain.



For example, we have already discussed the function

$$F:(x, y) \longmapsto \sqrt{1-(x^2+y^2)} \qquad ((x, y) \in \{(x, y): x^2+y^2 \leqslant 1\})$$

The domain has the circle $x^2 + y^2 = 1$ in the xy-plane as its boundary.

Suppose that we wish to find the overall maximum (or minimum) value of the images of a function F with domain, A, a subset of $R \times R$. The points where the tangent plane is parallel to the xy-plane, on the surface defined by $z = F(x, y)$, are clearly going to be of interest. This leads us to our next definition.

Definition 1

If $F_1'(a, b) = 0$ and $F_2'(a, b) = 0$, then (a, b) is called a stationary point of F.

Notice that since the tangent plane at a stationary point is parallel to the xy-plane, its equation is simply $z = F(a, b)$.

15.3.4 Local Maxima and Minima

15.3.4 Definitions

You may find the precise definitions of *local maximum* and *local minimum* a little hard to digest, so we give intuitive definitions first.

Intuitive Definition 1

If (a, b) is a point in the domain of F, and if $F(x, y) \leqslant F(a, b)$ for all (x, y) in the domain of F sufficiently close to (a, b), then we say that F has a local maximum at (a, b).

Intuitive Definition 2

If (a, b) is a point in the domain of F, and if $F(x, y) \geqslant F(a, b)$ for all (x, y) in the domain of F sufficiently close to (a, b), then we say that F has a local minimum at (a, b).

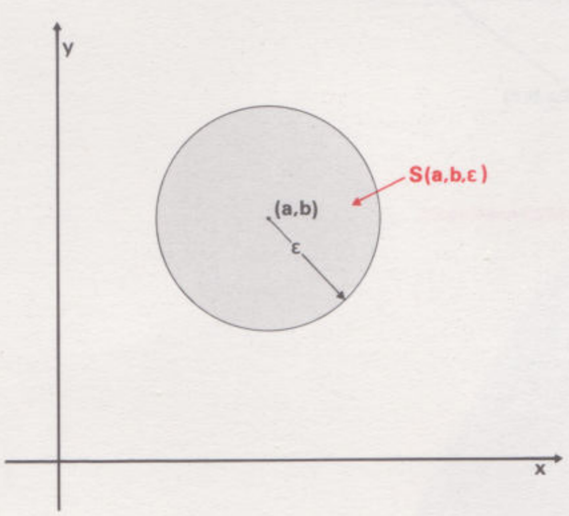
Speaking very roughly, if it rains on the surface, the puddles collect around the local minima, and the water runs away to the overall minimum as the puddles overflow.

Discussion

The difficulty with the above definitions is that they depend on the meaning of "sufficiently close", and it is this phrase which needs to be precisely defined.

If we use our approach to functions of one variable as a guide, then we need a "small" set in $R \times R$ where before we had a "small" interval, $[c - \varepsilon, c + \varepsilon]$ in R; the most suitable set in $R \times R$ is a circular disc.

We let $S(a, b, \varepsilon)$ denote the set $\{(x, y): (x - a)^2 + (y - b)^2 \leqslant \varepsilon^2\}$, which is a disc with centre at the point (a, b) and radius ε.

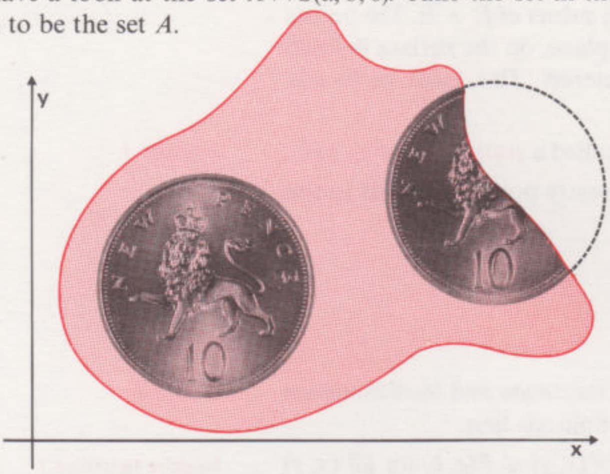


Let F be a function with domain $A \subseteq R \times R$. Then, following our definitions for functions of one variable, we make the following formal definitions.

If there is a positive number ε such that $F(x, y) \leqslant F(a, b)$ for all $(x, y) \in A \cap S(a, b, \varepsilon)$, then we say that F has a local maximum at (a, b). **Definition 1**

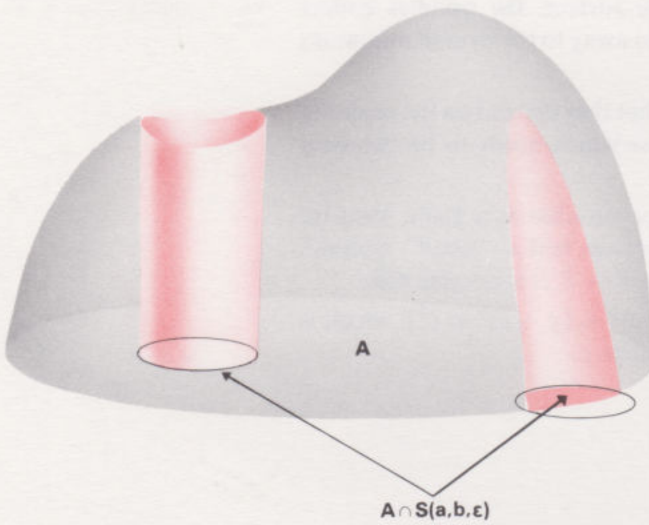
If there is a positive number ε such that $F(x, y) \geqslant F(a, b)$ for all $(x, y) \in A \cap S(a, b, \varepsilon)$, then we say that F has a local minimum at (a, b). **Definition 2**

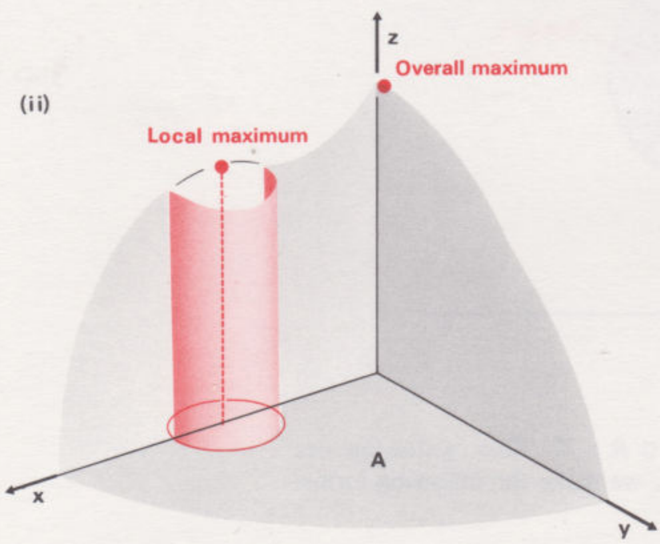
Let us have a look at the set $A \cap S(a, b, \varepsilon)$. Take the set in the following diagram to be the set A.



If we represent S by a 10p coin, then, placing the 10p down on the set, the part of A which is covered by the coin is the set $A \cap S(a, b, \varepsilon)$. The point (a, b) is of course the centre point of the coin, and ε is its radius.

Example 1 **Example 1**





■

Just as for functions of one variable, we have two major problems. A stationary point need not be a local maximum nor a local minimum. Stationary points of this kind are called saddle points (for obvious reasons).

Discussion
* *

Definition 3
* *

Horizontal tangent plane at a saddle point

In other words, saying that the tangent plane is horizontal guarantees neither a local maximum nor a local minimum.

The second big problem is that a local maximum, or indeed an overall maximum, can occur on the boundary of the domain, and similarly for local and overall minima. If we restrict ourselves to a search for stationary points, then we may miss points of this kind.

We shall attempt to overcome the first problem, but we shall not have time to find a way to overcome the second problem in this course. In general, the second problem is not so serious when applying the methods to physical situations, since the information from the situation itself often helps us to decide the nature of the point we are dealing with.

15.3.5 The Aqueduct Problem

15.3.5

Example
*

We are at last in a position to attempt to solve the engineer's problem of designing the aqueduct which we mentioned on page 33. We left him trying to decide on the choice of x and y which would give the overall maximum value of the function:

$$F:(x, y) \longmapsto (4 - 2x + y)\sqrt{x^2 - y^2}, \quad ((x, y) \in [0, 2] \times [0, 2] \text{ and } y \leqslant x).$$

It is important to notice that we do not need to draw a picture of the surface defined by the equation

$$z = (4 - 2x + y)\sqrt{x^2 - y^2},$$

in fact, we want to avoid this if possible.

We found in Exercise 15.3.2.2 that

$$F_1'(a, b) = -2\sqrt{a^2 - b^2} + \left(\frac{a}{\sqrt{a^2 - b^2}}\right)(4 - 2a + b)$$

and

$$F_2'(a, b) = \sqrt{a^2 - b^2} - \left(\frac{b}{\sqrt{a^2 - b^2}}\right)(4 - 2a + b).$$

The tangent plane is horizontal if $F'_1(a, b) = F'_2(a, b) = 0$, so we need to solve the pair of equations:

$$\left.\begin{aligned} -2\sqrt{a^2 - b^2} + \left(\frac{a}{\sqrt{a^2 - b^2}}\right)(4 - 2a + b) &= 0 \\ \sqrt{a^2 - b^2} - \left(\frac{b}{\sqrt{a^2 - b^2}}\right)(4 - 2a + b) \quad &= 0 \end{aligned}\right\}$$

which simplifies to

$$\left.\begin{aligned} \frac{a}{2}(4 - 2a + b) &= a^2 - b^2 \\ b(4 - 2a + b) &= a^2 - b^2 \end{aligned}\right\}.$$

These equations are satisfied simultaneously if $b = \dfrac{a}{2}$, and then, substituting for b in the first equation, we obtain $a = \frac{4}{3}$, and therefore $b = \frac{2}{3}$. (The other* solution, $a = b = 0$, clearly corresponds to a *minimum* value of $F(x, y)$, namely 0.) The corresponding value of the cross-sectional area is $\dfrac{4}{\sqrt{3}} \simeq 2.3$, which is obtained by replacing x by $\frac{4}{3}$ and y by $\frac{2}{3}$ in the expression for $F(x, y)$. (Remember that when the sides of the aqueduct are vertical the greatest value of the cross-sectional area is 2.)

The engineer would surely find this result very impressive. Just by doing a fairly simple calculation, he would be able to increase the flow of water by about 15% *at no extra cost in material.* He has a greater value for the area than before, so there is some cause for satisfaction. It would, of course, be even better if he could assure himself that he had found the greatest possible value. **Discussion** •

In this case we are fairly certain from the physical origins of the problem that the point where the tangent plane is horizontal will give a local maximum value of F. However, we cannot always be this certain, and it would help to have a technique which would distinguish between local maxima, local minima, and other points where the tangent plane is horizontal. This is our next task, and we shall begin by considering a very easy problem for which we know the answer in advance.

15.3.6. A Useful Technique

15.3.6

Example 1

Example 1

Find the overall minimum value of the function

$$F:(x, y) \longmapsto x^2 + 2y^2 \qquad ((x, y) \in R \times R). \qquad \blacksquare$$

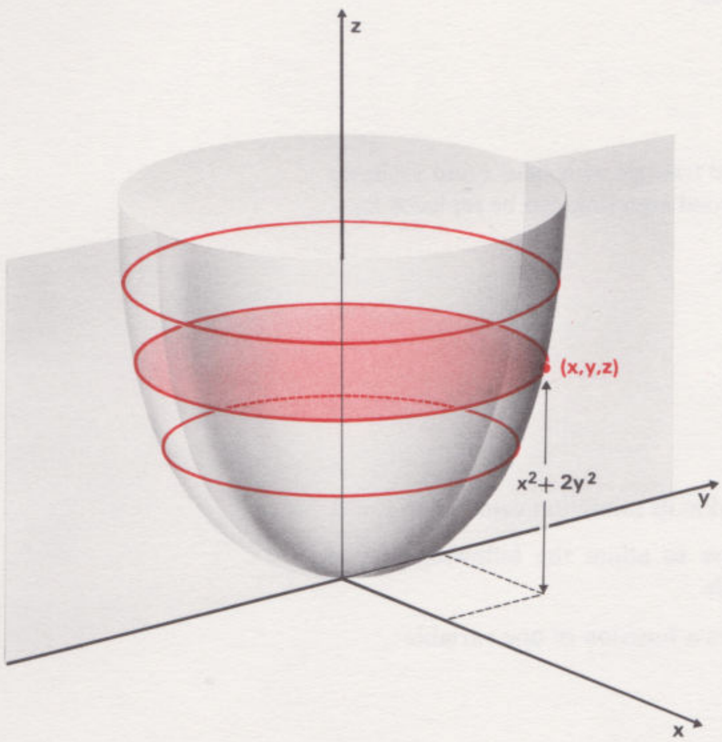
Solution of Example 1

Since both expressions on the right are positive or zero, the answer is obviously 0, and this occurs when $x = y = 0$.

Now let us use the above example to test a technique for classifying stationary points, which we can use when the answer isn't obvious. **Discussion** • •

* The solution $a = b = 4$ does not correspond to a point in the domain of F.

The surface $z = x^2 + 2y^2$ looks like this:



Notice that

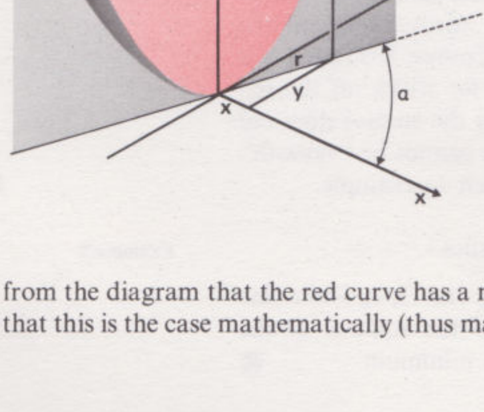
$$F'_1(a, b) = 2a \quad \text{and} \quad F'_2(a, b) = 4b,$$

and therefore the point corresponding to $a = b = 0$ is a stationary point, confirming what we already know.

The essential idea of our proposed technique is as follows: intersect the surface with the plane whose equation is

$$y = x \tan \alpha,$$

to give the curve shown in red.



It is obvious from the diagram that the red curve has a minimum at 0, but can we show that this is the case mathematically (thus making the diagram redundant)?

The red curve is determined geometrically as the intersection of the plane, $\{(x, y): y = x \tan \alpha\}$, and the surface, $\{(x, y): z = x^2 + 2y^2\}$, or, more briefly, it is determined by the equations

$$z = x^2 + 2y^2$$

$$y = x \tan \alpha.$$

If r is the hypotenuse of the right-angled triangle with sides x and y (shown on the previous diagram), then the second equation can be replaced by

$$x = r \cos \alpha \quad \text{and} \quad y = r \sin \alpha.$$

On the red curve we then have

$$\begin{aligned} z &= r^2(\cos^2 \alpha + 2 \sin^2 \alpha) \\ &= r^2(1 + \sin^2 \alpha), \end{aligned}$$

so that on the red curve, $z(= F(x, y))$ takes its minimum value when $r = 0$.

In more difficult cases we might have to adopt the following line of reasoning to achieve the required result.

The equation $z = r^2(1 + \sin^2 \alpha)$ defines a function of one variable:

$$\phi : r \longmapsto z \qquad (r \in R^+)$$

Differentiating, we obtain

$$\phi'(r) = 2r(1 + \sin^2 \alpha)$$

and

$$\phi''(r) = 2(1 + \sin^2 \alpha).$$

In particular, $\phi''(0) = 2(1 + \sin^2 \alpha)$, so that $\phi''(0) > 0$, and therefore the red curve has a *local minimum* at $r = 0$.

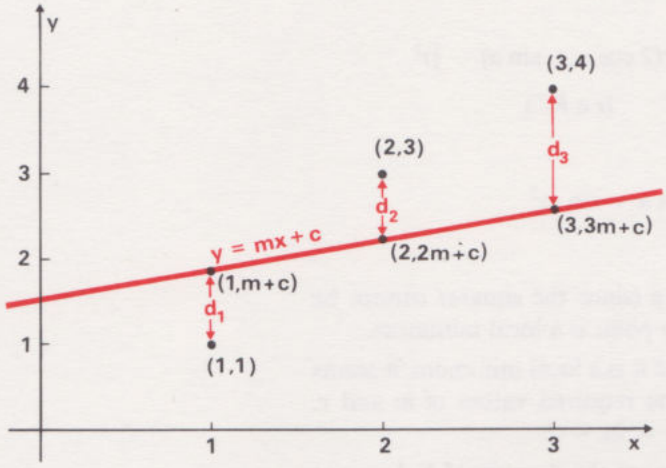
The essential point about the equation $\phi''(0) = 2(1 + \sin^2 \alpha)$ is that it shows that $\phi''(0) > 0$ *for all values of* α, so that all possible red curves obtained in this way have a local minimum at 0. This seems to show pretty conclusively that F has a local minimum at 0 (again confirming the result which we already know). ■

Roughly speaking, the technique can be summarized as follows. Slice through the point of interest on the surface with a vertical cut, revealing a curve like the red curve of our example. If all such curves have a local minimum at the point, then we would expect the surface to have a local minimum there. Our next example will again illustrate this idea.

This technique for classifying the stationary points will be adequate when the surface is "smooth", and it will certainly be adequate for all the problems which you will meet in the Foundation Course. However, it is an amazing fact that one can construct a surface for which all the red curves have a local minimum at the origin, *and yet* the surface does *not* have a local minimum at that point. Such a surface cannot be "smooth" in our sense, and you may like to try to think of such an example.

Example 2 (This example has applications in statistics.) **Example 2**

Given the three points with co-ordinates (1, 1), (2, 3) and (3, 4), find a line with equation $y = mx + c$, such that the sum of the squares of the "vertical" distances of the points from the line is a minimum. ■



Solution of Example 2

In the diagram, d_1, d_2 and d_3 are the vertical distances, and we want to minimize

$$d_1^2 + d_2^2 + d_3^2 = (m + c - 1)^2 + (2m + c - 3)^2 + (3m + c - 4)^2;$$

we can therefore define a function F with domain $R \times R$ by putting

$$F(m, c) = (m + c - 1)^2 + (2m + c - 3)^2 + (3m + c - 4)^2. \qquad \textbf{Equation (1)}$$

We then have

$$F'_1(m, c) = 2(m + c - 1) + 4(2m + c - 3) + 6(3m + c - 4)$$
$$= 28m + 12c - 38$$

and

$$F'_2(m, c) = 2(m + c - 1) + 2(2m + c - 3) + 2(3m + c - 4)$$
$$= 12m + 6c - 16.$$

The values of m and c for which $F'_1(m, c) = F'_2(m, c) = 0$ are determined by the equations

$$14m + 6c - 19 = 0$$
$$12m + 6c - 16 = 0$$

from which we deduce that

$$m = \tfrac{3}{2} \quad \text{and} \quad c = -\tfrac{1}{3}.$$

These values of m and c determine a stationary point of F, but is this stationary point the overall minimum? Could we show that the stationary point is even a local minimum? Let us try the method of our previous example.

Our variables are m and c instead of x and y, and so we intersect the unknown surface by planes perpendicular to the mc-plane and through the point $(\frac{3}{2}, -\frac{1}{3})$. A typical plane has equation

$$c + \tfrac{1}{3} = (m - \tfrac{3}{2}) \tan \alpha$$

and just as before we can replace this equation by

$$m = \tfrac{3}{2} + r \cos \alpha \qquad c = -\tfrac{1}{3} + r \sin \alpha$$

where r is now the distance of the point with co-ordinates (m, c) from the point $(\frac{3}{2}, -\frac{1}{3})$. Substituting these expressions for m and c into Equation (1),

we obtain the function ϕ defined by

$$\phi(r) = (r(\cos\alpha + \sin\alpha) + \tfrac{1}{6})^2 + (r(2\cos\alpha + \sin\alpha) - \tfrac{1}{3})^2 + (r(3\cos\alpha + \sin\alpha) + \tfrac{1}{6})^2 \qquad (r \in R^+).$$

Differentiating twice

$$\phi''(0) = 2((\cos\alpha + \sin\alpha)^2 + (2\cos\alpha + \sin\alpha)^2 + (3\cos\alpha + \sin\alpha)^2)$$

and therefore $\phi''(0) > 0$ for *all* values of α (since the squares cannot be zero simultaneously); hence the stationary point is a local minimum.

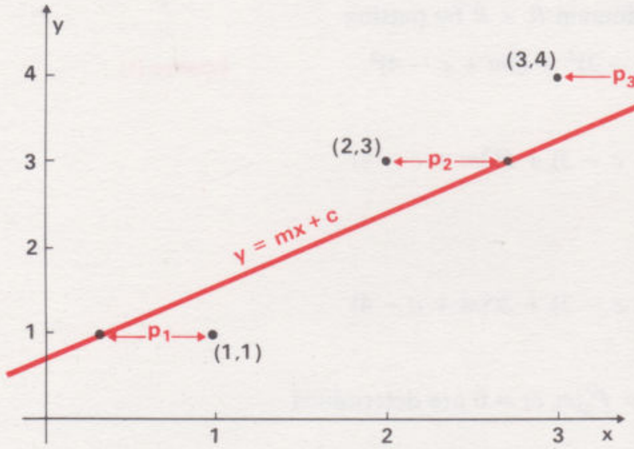
Since there is only one stationary point and it is a local minimum, it seems very likely that we have indeed found the required values of m and c. So the equation of the required line is $6y = 9x - 2$.

The difficulty with points on the boundary of the domain of F does not occur in this case, because the domain is the whole set $R \times R$, and there are no boundary points; but to be safe we ought really to find the images of the function when r is very large. For the moment we shall avoid this difficulty too. ■

Exercise 1

Exercise 1
(4 minutes)

Find the equation of the red line which gives the minimum value of $p_1^2 + p_2^2 + p_3^2$.

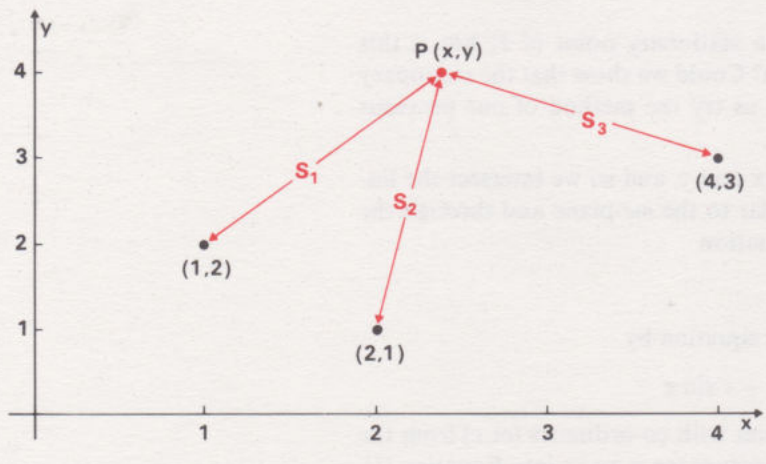


■

Exercise 2

Exercise 2
(3 minutes)

Find the point P for which the sum $S_1^2 + S_2^2 + S_3^2$ is a minimum.



■

15.4 SUMMARY AND CONCLUSIONS

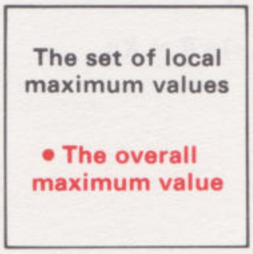
15.4

15.4.1 Summary

15.4.1

Summary

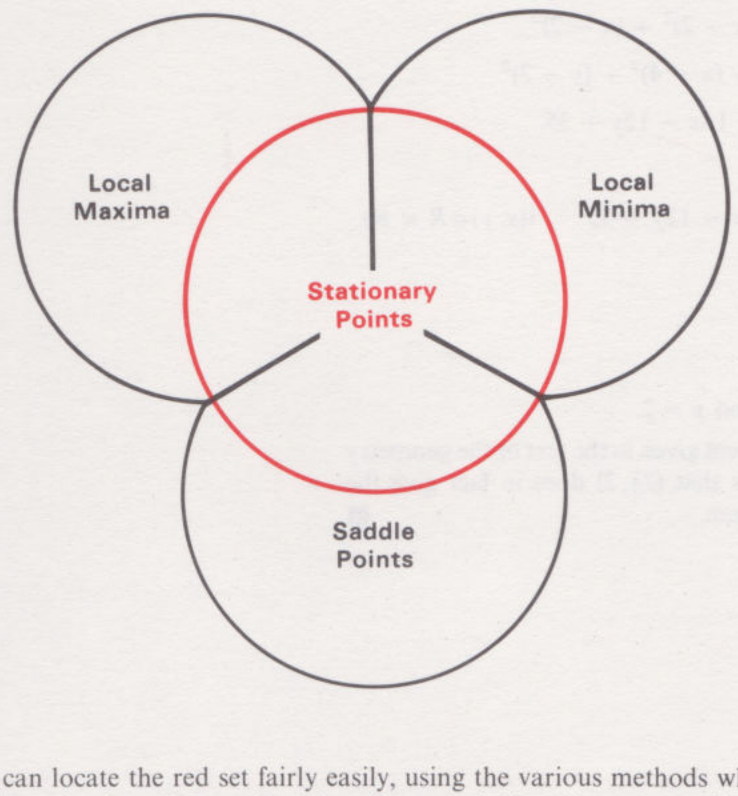
Our original problem was to find the overall maximum value attained by the images of a function *f*. This problem is difficult, but we can go some way towards a solution if we can locate the local maxima. We know that the overall maximum value belongs to the set of local maximum values; it is the greatest element of that set.



We hope to locate the local maxima by first locating the stationary points and then examining their nature, but there are two complications:

(i) local maxima at the boundary of the domain need not be stationary points;

(ii) stationary points can also be local minima or saddle points.



We can locate the red set fairly easily, using the various methods which we have developed. For a function of one variable with an interval $[a, b]$ as its domain there are are only two boundary points, and we can use the strategy which we developed in section 15.1. For a function of two variables the situation is more difficult, for in this case there are, in general, an infinite number of boundary points. Our previous strategy will not work in this case, because we cannot test an infinite number of values to find the largest. But often the physical origins of the problem give us useful information which allows us to resolve the problem.

Solution 15.3.6.1

Solution 15.3.6.1

Writing the equation of the line in the form $y = mx + c$ leads to some untidy algebra. In this exercise it is convenient to take $x = my + c$ as the equation of the line; we then have:

$$\begin{aligned} p_1^2 + p_2^2 + p_3^2 &= (1-(m+c))^2 + (2-(3m+c))^2 + (3-(4m+c))^2 \\ &= 26m^2 + 16mc + 3c^2 - 38m - 12c + 14. \end{aligned}$$

So F is the function defined by

$$F:(m,c) \longmapsto 26m^2 + 16mc + 3c^2 - 38m - 12c + 14 \qquad ((m,c) \in R \times R).$$

Thus, for a stationary point, we have

$$F_1'(m,c) = 52m + 16c - 38 = 0$$

$$F_2'(m,c) = 16m + 6c - 12 = 0.$$

The solution of this pair of equations is

$$m = \tfrac{9}{14}, \qquad c = \tfrac{2}{7}.$$

An argument similar to that given in the text would show that these values of m and c do give a local minimum. Thus the equation of the line is

$$14x = 9y + 4.$$

■

Solution 15.3.6.2

Solution 15.3.6.2

The sum

$$\begin{aligned} S_1^2 + S_2^2 + S_3^2 &= (x-1)^2 + (y-2)^2 + (x-2)^2 \\ &\quad + (y-1)^2 + (x-4)^2 + (y-3)^2 \\ &= 3x^2 + 3y^2 - 14x - 12y + 35. \end{aligned}$$

So we let F be the function

$$F:(x,y) \longmapsto 3x^2 + 3y^2 - 14x - 12y + 35 \qquad ((x,y) \in R \times R).$$

We have

$$F_1'(x,y) = 6x - 14$$

$$F_2'(x,y) = 6y - 12.$$

Thus for a stationary point, $x = 2\frac{1}{3}$ and $y = 2$.

Once again we can use either the argument given in the text or the geometry of the situation to convince ourselves that $(2\frac{1}{3}, 2)$ does in fact give the point P for which the sum is a minimum.

■

15.4.2 In Conclusion

15.4.2

Discussion

Our concern in this unit has been almost entirely with the optimization of functions, and our arguments have been based wholly on plausible geometric notions. As we have mentioned several times, things can go wildly wrong.

Our reasoning has been almost entirely heuristic, and it is worth quoting Polya on this point (see *Polya* page 113):

> "Heuristic reasoning is reasoning not regarded as final and strict but as provisional and plausible only, whose purpose is to discover the solution of the present problem. We are often obliged to use heuristic reasoning. We shall attain complete certainty when we shall have obtained the complete solution, but before obtaining certainty we must often be satisfied with a more or less plausible guess. We may need the provisional before we attain the final. We need heuristic reasoning when we construct a strict proof as we need scaffolding when we erect a building... Heuristic reasoning is good in itself. What is bad is to mix up heuristic reasoning with rigorous proof. What is worse is to sell heuristic reasoning for rigorous proof."

You may have noticed that we have been careful to say that our results are "likely" or "plausible", but never "certain". The main point is that our geometric intuition might lead us astray. We could hope to overcome this by giving a rigorous definition of what we mean by a surface, and in particular a "smooth" or "well-behaved" surface. But would such an approach be helpful if we wished to extend our methods to functions of more than two variables?

The solution to our difficulties is in fact to abandon the geometric notions entirely, and to base our ideas on a few non-geometric axioms. This will be one of the tasks of the analysis section of a later course. We shall, as it were, use the "geometric scaffolding" only as a guide when the building starts in earnest.

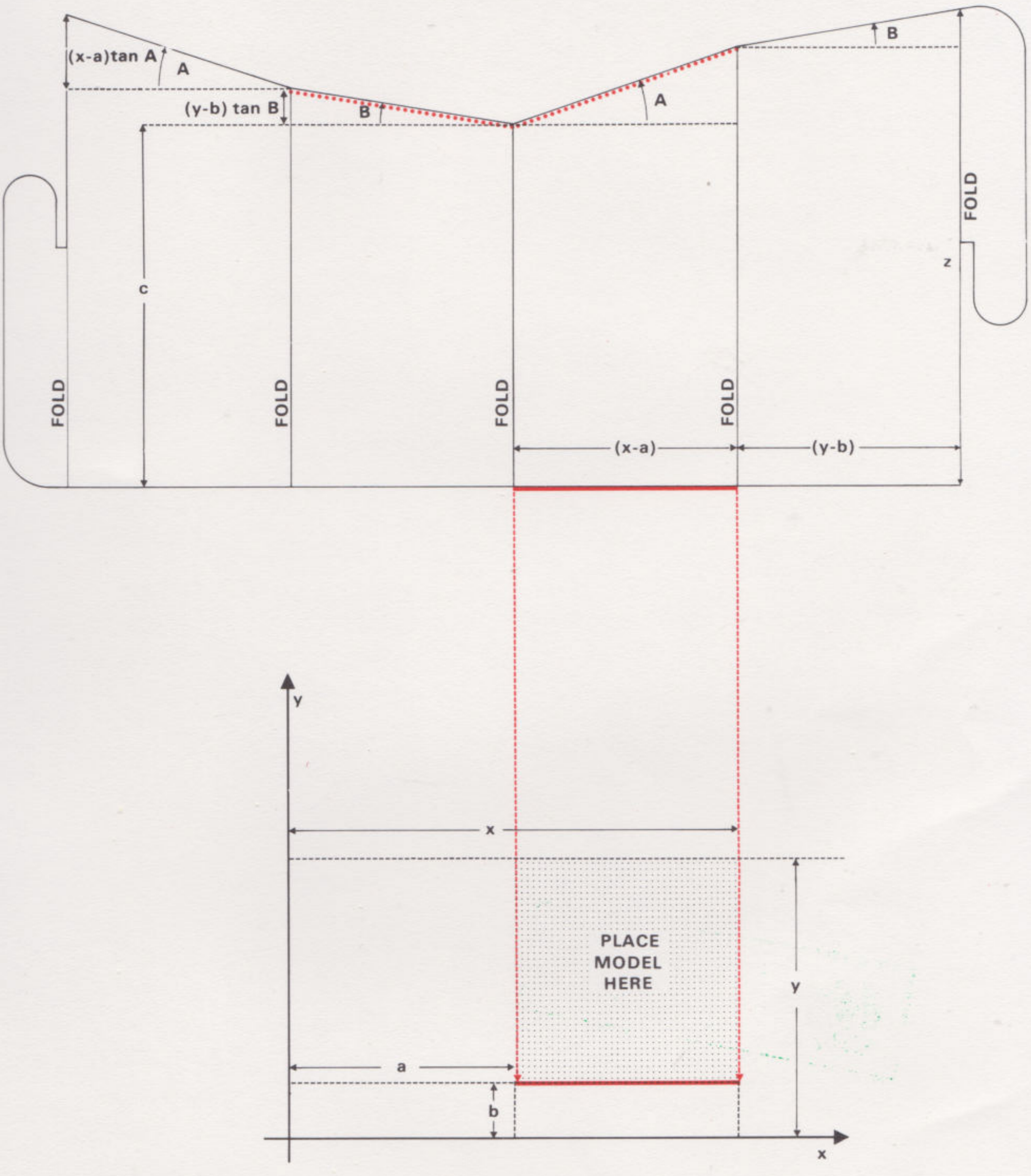
Don't be misled into thinking that the geometric reasoning has been a waste of time, for now we have a very clear idea of the results which we would like to prove formally. In the future we may well allow our geometric reasoning to lead the way, but we hope that it will be closely followed by rigorous analysis.

We would like to leave you with one or two thoughts. Could we approximate functions of two (or more) variables using an extension of Taylor's series for functions of one variable? Is there any way of deciding if a turning point of a function of two variables is a saddle-point? In the problem of the aqueduct, is there some other way of shaping the metal plates which would give an even greater cross-sectional area?

Acknowledgement

Grateful acknowledgement is made to the following source for the illustration used in this unit:
Page 42, the map — based upon the Ordnance Survey Sheet 111, with the sanction of the Controller of H.M. Stationery Office. Crown Copyright reserved.

Unit No.		Title of Text
1		Functions
2		Errors and Accuracy
3		Operations and Morphisms
4		Finite Differences
5	NO TEXT	
6		Inequalities
7		Sequences and Limits I
8		Computing I
9		Integration I
10	NO TEXT	
11		Logic I — Boolean Algebra
12		Differentiation I
13		Integration II
14		Sequences and Limits II
15		Differentiation II
16		Probability and Statistics I
17		Logic II — Proof
18		Probability and Statistics II
19		Relations
20		Computing II
21		Probability and Statistics III
22		Linear Algebra I
23		Linear Algebra II
24		Differential Equations I
25	NO TEXT	
26		Linear Algebra III
27		Complex Numbers I
28		Linear Algebra IV
29		Complex Numbers II
30		Groups I
31		Differential Equations II
32	NO TEXT	
33		Groups II
34		Number Systems
35		Topology
36		Mathematical Structures


(x-a)tan A
A
(y-b) tan B
B
A
B
c
FOLD
FOLD
FOLD
FOLD
FOLD
z
(x-a)
(y-b)
y
x
PLACE
MODEL
HERE
y
a
b
x